

20/20

```

1441 TTG GTG AAC AGG CGA CCA TGC TTT TCA GCT CTG GAA GTC GAT GAA ACA TAC GTT CCC AAA 1500
481 L V N R R P C F S A L E V D E T Y V P K 500

1501 GAG TTT AAT GCT GAA ACA TTC ACC TTC CAT GCA GAT ATA TGC ACA CTT TCT GAG AAG GAG 1560
501 E F N A E T F T F H A D I C T L S E K E 520

1561 AGA CAA ATC AAG AAA CAA ACT GCA CTT GTT GAG CTT GTG AAA CAC AAG CCC AAG GCA ACA 1620
521 R Q I K K Q T A L V E L V K H K P K A T 540

1621 AAA GAG CAA CTG AAA GCT GTT ATG GAT GAT TTC GCA GCT TTT GTA GAG AAG TGC TGC AAG 1680
541 K E Q L K A V M D D F A A F V E K C C K 560

1681 GCT GAC GAT AAG GAG ACC TGC TTT GCC GAG GAG GGT AAA AAA CTT GTT GCT GCA AGT CAA 1740
561 A D D K E T C F A E E G K K L V A A S Q 580

1741 GCT GCC TTA GGC TTA TAA CAT CTA CAT TTA AAA GCA TCT CAG 1782
581 A A L G L * 585

```

Figure 15D

## SEQUENCE LISTING

<110> Human Genome Science, Inc.  
Principia Pharmaceutical Corporation

<120> Albumin Fusion Proteins

<130> PF543PCT

<140> Unassigned

<141> 2001-04-12

<150> 60/229,358

<151> 2000-04-12

<150> 60/256,931

<151> 2000-12-21

<150> 60/199,384

<151> 2000-04-25

<160> 36

<170> PatentIn Ver. 2.1

<210> 1

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<221> primer\_bind

<223> primer useful to clone human growth hormone cDNA

<400> 1

cccaagaatt cccttatcca ggc

23

<210> 2

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<221> primer\_bind

<223> primer useful to clone human growth hormone cDNA

<400> 2

gggaagctta gaagccacag gatccctcca cag

33

<210> 3

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_structure

<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 3

gataaagatt cccaac

16

<210> 4

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_structure

<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 4

aattggtggg aatcttt

17

<210> 5

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_structure

<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 5

ttaggcttat tccaac

17

<210> 6

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<221> misc\_structure

<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 6

aattggtggg aataagcc

18

<210> 7

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<221> SITE

<222> 1)..(19)

<223> invertase leader sequence

<220>

<221> SITE  
 <222> 20)..(24)  
 <223> first 5 amino acids of mature human serum albumin  
  
 <400> 7  
 Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys  
       1                  5                  10                  15  
  
 Ile Ser Ala Asp Ala His Lys Ser  
                   20

<210> 8  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_structure  
 <223> synthetic oligonucleotide used to join DNA  
 fragments with non-cohesive ends.

<400> 8  
 gagatgcaca cctgagtgag g 21

<210> 9  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_structure  
 <223> synthetic oligonucleotide used to join DNA  
 fragments with non-cohesive ends.

<400> 9  
 gatcctgtgg cttcgatgca cacaaga 27

<210> 10  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_structure  
 <223> synthetic oligonucleotide used to join DNA  
 fragments with non-cohesive ends.

<400> 10  
 ctcttggtgtg catcgaagcc acag 24

<210> 11  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>



<221> misc\_structure  
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 11  
tgtggaagag cctcagaatt tattcccaac 30

<210> 12  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_structure  
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 12  
aattgttggg aataaattct gaggtcttc c 31

<210> 13  
<211> 47  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_structure  
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 13  
ttaggcttag gtggcggtgg atccggcggt ggtggatctt tcccaac 47

<210> 14  
<211> 48  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_structure  
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 14  
aattgttggg aaagatccac caccgccgga tccaccgcca cctaagcc 48

<210> 15  
<211> 62  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> misc\_structure  
<223> synthetic oligonucleotide used to join DNA fragments with non-cohesive ends.

<400> 15  
 ttaggcttag gcggtggtgg atctggtggc ggcggatctg gtggcgggtgg atccttccca 60  
 ac 62

<210> 16  
 <211> 63  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> misc\_structure  
 <223> synthetic oligonucleotide used to join DNA  
 fragments with non-cohesive ends.

<400> 16  
 aattgttggg aaggatccac cgccaccaga tccgccgcca ccagatccac caccgcctaa 60  
 gcc 63

<210> 17  
 <211> 1782  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (1)..(1755)

<400> 17  
 gat gca cac aag agt gag gtt gct cat cgg ttt aaa gat ttg gga gaa 48  
 Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu  
 1 5 10 15  
 gaa aat ttc aaa gcc ttg gtg ttg att gcc ttt gct cag tat ctt cag 96  
 Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln  
 20 25 30  
 cag tgt cca ttt gaa gat cat gta aaa tta gtg aat gaa gta act gaa 144  
 Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu  
 35 40 45  
 ttt gca aaa aca tgt gtt gct gat gag tca gct gaa aat tgt gac aaa 192  
 Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys  
 50 55 60  
 tca ctt cat acc ctt ttt gga gac aaa tta tgc aca gtt gca act ctt 240  
 Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu  
 65 70 75 80  
 cgt gaa acc tat ggt gaa atg gct gac tgc tgt gca aaa caa gaa cct 288  
 Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro  
 85 90 95  
 gag aga aat gaa tgc ttc ttg caa cac aaa gat gac aac cca aac ctc 336  
 Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu  
 100 105 110  
 ccc cga ttg gtg aga cca gag gtt gat gtg atg tgc act gct ttt cat 384

Pro	Arg	Leu	Val	Arg	Pro	Glu	Val	Asp	Val	Met	Cys	Thr	Ala	Phe	His		
		115					120					125					
gac	aat	gaa	gag	aca	ttt	ttg	aaa	aaa	tac	tta	tat	gaa	att	gcc	aga	432	
Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys	Lys	Tyr	Leu	Tyr	Glu	Ile	Ala	Arg		
	130					135					140						
aga	cat	cct	tac	ttt	tat	gcc	ccg	gaa	ctc	ctt	ttc	ttt	gct	aaa	agg	480	
Arg	His	Pro	Tyr	Phe	Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe	Ala	Lys	Arg		
145					150					155					160		
tat	aaa	gct	gct	ttt	aca	gaa	tgt	tgc	caa	gct	gct	gat	aaa	gct	gcc	528	
Tyr	Lys	Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala		
				165					170					175			
tgc	ctg	ttg	cca	aag	ctc	gat	gaa	ctt	cgg	gat	gaa	ggg	aag	gct	tcg	576	
Cys	Leu	Leu	Pro	Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser		
			180					185					190				
tct	gcc	aaa	cag	aga	ctc	aaa	tgt	gcc	agt	ctc	caa	aaa	ttt	gga	gaa	624	
Ser	Ala	Lys	Gln	Arg	Leu	Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu		
		195					200					205					
aga	gct	ttc	aaa	gca	tgg	gca	gtg	gct	cgc	ctg	agc	cag	aga	ttt	ccc	672	
Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val	Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro		
	210					215					220						
aaa	gct	gag	ttt	gca	gaa	gtt	tcc	aag	tta	gtg	aca	gat	ctt	acc	aaa	720	
Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu	Val	Thr	Asp	Leu	Thr	Lys		
225					230					235				240			
gtc	cac	acg	gaa	tgc	tgc	cat	gga	gat	ctg	ctt	gaa	tgt	gct	gat	gac	768	
Val	His	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu	Cys	Ala	Asp	Asp		
			245						250					255			
agg	gcg	gac	ctt	gcc	aag	tat	atc	tgt	gaa	aat	cag	gat	tcg	atc	tcc	816	
Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser	Ile	Ser		
			260					265					270				
agt	aaa	ctg	aag	gaa	tgc	tgt	gaa	aaa	cct	ctg	ttg	gaa	aaa	tcc	cac	864	
Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His		
		275					280					285					
tgc	att	gcc	gaa	gtg	gaa	aat	gat	gag	atg	cct	gct	gac	ttg	cct	tca	912	
Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser		
	290					295					300						
tta	gct	gct	gat	ttt	gtt	gaa	agt	aag	gat	gtt	tgc	aaa	aac	tat	gct	960	
Leu	Ala	Ala	Asp	Phe	Val	Glu	Ser	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala		
305					310					315				320			
gag	gca	aag	gat	gtc	ttc	ctg	ggc	atg	ttt	ttg	tat	gaa	tat	gca	aga	1008	
Glu	Ala	Lys	Asp	Val	Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg		
				325					330					335			
agg	cat	cct	gat	tac	tct	gtc	gtg	ctg	ctg	ctg	aga	ctt	gcc	aag	aca	1056	
Arg	His	Pro	Asp	Tyr	Ser	Val	Val	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr		
			340					345					350				
tat	gaa	acc	act	cta	gag	aag	tgc	tgt	gcc	gct	gca	gat	cct	cat	gaa	1104	

Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys	Cys	Ala	Ala	Ala	Asp	Pro	His	Glu		
		355					360					365					
tgc	tat	gcc	aaa	gtg	ttc	gat	gaa	ttt	aaa	cct	ctt	gtg	gaa	gag	cct	1152	
Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	Phe	Lys	Pro	Leu	Val	Glu	Glu	Pro		
	370					375					380						
cag	aat	tta	atc	aaa	caa	aac	tgt	gag	ctt	ttt	gag	cag	ctt	gga	gag	1200	
Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu	Gln	Leu	Gly	Glu		
385					390					395				400			
tac	aaa	ttc	cag	aat	gcg	cta	tta	gtt	cgt	tac	acc	aag	aaa	gta	ccc	1248	
Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys	Lys	Val	Pro		
				405					410					415			
caa	gtg	tca	act	cca	act	ctt	gta	gag	gtc	tca	aga	aac	cta	gga	aaa	1296	
Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys		
			420					425					430				
gtg	ggc	agc	aaa	tgt	tgt	aaa	cat	cct	gaa	gca	aaa	aga	atg	ccc	tgt	1344	
Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro	Glu	Ala	Lys	Arg	Met	Pro	Cys		
		435					440					445					
gca	gaa	gac	tat	cta	tcc	gtg	gtc	ctg	aac	cag	tta	tgt	gtg	ttg	cat	1392	
Ala	Glu	Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu	Cys	Val	Leu	His		
	450					455					460						
gag	aaa	acg	cca	gta	agt	gac	aga	gtc	aca	aaa	tgc	tgc	aca	gag	tcc	1440	
Glu	Lys	Thr	Pro	Val	Ser	Asp	Arg	Val	Thr	Lys	Cys	Cys	Thr	Glu	Ser		
465					470					475				480			
ttg	gtg	aac	agg	cga	cca	tgc	ttt	tca	gct	ctg	gaa	gtc	gat	gaa	aca	1488	
Leu	Val	Asn	Arg	Arg	Pro	Cys	Phe	Ser	Ala	Leu	Glu	Val	Asp	Glu	Thr		
				485					490					495			
tac	gtt	ccc	aaa	gag	ttt	aat	gct	gaa	aca	ttc	acc	ttc	cat	gca	gat	1536	
Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala	Glu	Thr	Phe	Thr	Phe	His	Ala	Asp		
			500					505					510				
ata	tgc	aca	ctt	tct	gag	aag	gag	aga	caa	atc	aag	aaa	caa	act	gca	1584	
Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg	Gln	Ile	Lys	Lys	Gln	Thr	Ala		
		515					520					525					
ctt	gtt	gag	ctt	gtg	aaa	cac	aag	ccc	aag	gca	aca	aaa	gag	caa	ctg	1632	
Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro	Lys	Ala	Thr	Lys	Glu	Gln	Leu		
	530					535					540						
aaa	gct	gtt	atg	gat	gat	ttc	gca	gct	ttt	gta	gag	aag	tgc	tgc	aag	1680	
Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys	Cys	Cys	Lys		
545					550					555				560			
gct	gac	gat	aag	gag	acc	tgc	ttt	gcc	gag	gag	ggt	aaa	aaa	ctt	gtt	1728	
Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys	Lys	Leu	Val		
				565					570					575			
gct	gca	agt	caa	gct	gcc	tta	ggc	tta	taacatctac	attttaaagc	atctcag	1782					
Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu									
			580				585										

<210> 18  
 <211> 585  
 <212> PRT  
 <213> Homo Sapiens

<400> 18  
 Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu  
     1                    5                    10                    15  
 Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln  
                     20                    25                    30  
 Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu  
                     35                    40                    45  
 Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys  
                     50                    55                    60  
 Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu  
     65                    70                    75                    80  
 Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro  
                     85                    90                    95  
 Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu  
                     100                    105                    110  
 Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His  
                     115                    120                    125  
 Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg  
     130                    135                    140  
 Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg  
     145                    150                    155                    160  
 Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala  
                     165                    170                    175  
 Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser  
                     180                    185                    190  
 Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu  
                     195                    200                    205  
 Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro  
     210                    215                    220  
 Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys  
     225                    230                    235                    240  
 Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp  
                     245                    250                    255  
 Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser  
                     260                    265                    270  
 Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His  
     275                    280                    285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser  
 290 295 300  
 Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala  
 305 310 315 320  
 Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg  
 325 330 335  
 Arg His Pro Asp Tyr Ser Val Val Leu Leu Arg Leu Ala Lys Thr  
 340 345 350  
 Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu  
 355 360 365  
 Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro  
 370 375 380  
 Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu  
 385 390 395 400  
 Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro  
 405 410 415  
 Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys  
 420 425 430  
 Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys  
 435 440 445  
 Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His  
 450 455 460  
 Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser  
 465 470 475 480  
 Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr  
 485 490 495  
 Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp  
 500 505 510  
 Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala  
 515 520 525  
 Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu  
 530 535 540  
 Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys  
 545 550 555 560  
 Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val  
 565 570 575  
 Ala Ala Ser Gln Ala Ala Leu Gly Leu  
 580 585

&lt;210&gt; 19

<211> 57  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> primer used to generate XhoI and ClaI  
site in pPPC0006

<400> 19  
gcctcgagaa aagagatgca cacaagagtg aggttgctca tcgatttaaa gatttgg 57

<210> 20  
<211> 58  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> primer used in generation XhoI and ClaI  
site in pPPC0006

<400> 20  
aatcgatgag caacctcact cttgtgtgca tctcttttct cgaggctcct ggaataag 58

<210> 21  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> primer used in generation XhoI and ClaI  
site in pPPC0006

<400> 21  
tacaaactta agagtccaat tagc 24

<210> 22  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> primer used in generation XhoI and ClaI  
site in pPPC0006

<400> 22  
cacttctcta gagtgggttc atatgtctt 29

<210> 23  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> Misc\_Structure  
<223> Synthetic oligonucleotide used to alter restriction sites in pPPC0007

<400> 23  
aagctgcctt aggcttataa taaggcgcgc cggccggccg tttaaactaa gcttaattct 60

<210> 24  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> Misc\_Structure  
<223> Synthetic oligonucleotide used to alter restriction sites in pPPC0007

<400> 24  
agaattaagc ttagttttaa cggccggccg gcgcgcctta ttataagcct aaggcagctt 60

<210> 25  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> forward primer useful for generation of albumin fusion protein in which the albumin moiety is N-terminal of the Therapeutic Protein

<220>  
<221> misc\_feature  
<222> (18)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (19)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (20)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (21)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature



<222> (23)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (24)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (25)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (26)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (27)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (28)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (29)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (30)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (31)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (32)  
<223> n equals a,t,g, or c

<400> 25  
aagctgcctt aggcttannn nnnnnnnnnn nn

32

<210> 26  
<211> 51  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> reverse primer useful for generation of albumin  
fusion protein in which the albumin moiety is N-terminal

of the Therapeutic Protein

<220>  
<221> misc\_feature  
<222> (37)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (38)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (39)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (40)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (41)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (42)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (43)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (44)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (45)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (46)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (47)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (48)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (49)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (50)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (51)

<223> n equals a,t,g, or c

<400> 26

gcgcgcgttt aaacggccgg ccggcgcgcc ttattannnn nnnnnnnnnn n

51

<210> 27

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> forward primer useful for generation of albumin fusion protein in which the albumin moiety is c-terminal of the Therapeutic Protein

<220>

<221> misc\_feature

<222> (19)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (20)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (21)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (22)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (23)

<223> n equals a,t,g, or c

<220>

<221> misc\_feature

<222> (24)

<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (25)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (26)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (27)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (28)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (29)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (30)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (31)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (32)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (33)  
<223> n equals a,t,g, or c

<400> 27  
aggagcgtcg acaaaagann nnnnnnnnnnn nnn

33

<210> 28  
<211> 52  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> reverse primer useful for generation of albumin  
fusion protein in which the albumin moiety is c-terminal of  
the Therapeutic Protein

<220>  
<221> misc\_feature  
<222> (38)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (39)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (40)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (41)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (42)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (43)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (44)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (45)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (46)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (47)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (48)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (49)  
<223> n equals a,t,g, or c

<220>  
 <221> misc\_feature  
 <222> (50)  
 <223> n equals a,t,g, or c

<220>  
 <221> misc\_feature  
 <222> (51)  
 <223> n equals a,t,g, or c

<220>  
 <221> misc\_feature  
 <222> (52)  
 <223> n equals a,t,g, or c

<400> 28  
 ctttaaatacg atgagcaacc tcactcttgt gtgcacnnnn nnnnnnnnnn nn 52

<210> 29  
 <211> 24  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <221> signal  
 <223> signal peptide of natural human serum albumin protein

<400> 29  
 Met Lys Trp Val Ser Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala  
           1                  5                  10                  15

Tyr Ser Arg Ser Leu Asp Lys Arg  
                   20

<210> 30  
 <211> 114  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> primer\_bind  
 <223> forward primer useful for generation of PC4:HSA  
 albumin fusion VECTOR

<220>  
 <221> misc\_feature  
 <222> (5)..(10)  
 <223> BamHI restriction site

<220>  
 <221> misc\_feature  
 <222> (11)..(16)  
 <223> Hind III restriction site

<220>  
 <221> misc\_feature  
 <222> (17)..(27)  
 <223> Kozak sequence

```

<220>
<221> misc_feature
<222> (25)..(97)
<223> cds natural signal sequence of human serum albumin

<220>
<221> misc_feature
<222> (75)..(81)
<223> XhoI restriction site

<220>
<221> misc_feature
<222> (98)..(114)
<223> cds first six amino acids of human serum albumin

<400> 30
tcagggatcc aagcttccgc caccatgaag tgggtaacct ttatttccct tctttttctc 60

tttagctcgg cttactcgag ggggtgtgtt cgtcgagatg cacacaagag tgag      114

<210> 31
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<221> primer_bind
<223> reverse primer useful for generation of
PC4:HSA albumin fusion VECTOR

<220>
<221> misc_feature
<222> (6)..(11)
<223> Asp718 restriction site

<220>
<221> misc_feature
<222> (12)..(17)
<223> EcoRI restriction site

<220>
<221> misc_feature
<222> (15)..(17)
<223> reverse complement of stop codon

<220>
<221> misc_feature
<222> (18)..(25)
<223> AscI restriction site

<220>
<221> misc_feature
<222> (18)..(43)
<223> reverse complement of DNA sequence encoding last 9 amino acids

<400> 31
gcagcggtag cgaattcggc ggcgcttata agcctaaggc agc      43

```

<210> 32  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> forward primer useful for inserting Therapeutic protein into pC4:HSA vector

<220>  
<221> misc\_feature  
<222> (29)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (30)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (31)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (32)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (33)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (34)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (35)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (36)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (37)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (38)  
<223> n equals a,t,g, or c



<220>  
<221> misc\_feature  
<222> (39)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (40)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (41)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (42)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (43)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (44)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (45)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (46)  
<223> n equals a,t,g, or c

<400> 32  
ccgccgctcg aggggtgtgt ttcgtcgann nnnnnnnnnn nnnnnn

<210> 33  
<211> 55  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> primer\_bind  
<223> reverse primer useful for inserting Therapeutic protein into pC4:HSA vector

<220>  
<221> misc\_feature  
<222> (38)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature

46

<222> (39)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (40)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (41)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (42)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (43)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (44)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (45)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (46)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (47)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (48)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (49)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature  
<222> (50)  
<223> n equals a,t,g, or c

<220>  
<221> misc\_feature

```
<222> (51)
<223> n equals a,t,g, or c
```

```
<220>  
<221> misc_feature  
<222> (52)  
<223> n equals a,t,g, or c
```

```
<220>  
<221> misc_feature  
<222> (53)  
<223> n equals a,t,g, or c
```

```
<220>  
<221> misc_feature  
<222> (54)  
<223> n equals a,t,g, or c
```

```
<220>  
<221> misc_feature  
<222> (55)  
<223> n equals a,t,g, or c
```

```
<400> 33
agtcccatcg atgagcaacc tcactcttgt gtgcatacnnn nnnnnnnnnn nnnnnn      55
```

```
<210> 34
<211> 17
<212> PRT
<213> Artificial Sequence
```

```
<220>
<221> signal
<223> Stanniocalcin signal peptide
```

```
<400> 34
Met Leu Gln Asn Ser Ala Val Leu Leu Leu Leu Val Ile Ser Ala Ser
   1                   5               10              15
```

Ala

```
<210> 35
<211> 22
<212> PRT
<213> Artificial Sequence
```

```
<220>
<221> signal
<223> Synthetic signal peptide
```

<400> 35  
Met Pro Thr Trp Ala Trp Trp Leu Phe Leu Val Leu Leu Leu Ala Leu  
1 5 10 15

Trp Ala Pro Ala Arg Gly  
20

<210>	36
<211>	402
<212>	DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 36

```
gctccaactt cttctttctac taagaagact caattgcaat tggaacactt gttggttgac 60
ttgcaaataga tcttaaaccg tataaacaac tataaaaacc caaagttgac tagaatgttg 120
actttcaagt tctacatgcc aaagaaagct actgaattga agcacttgca atgtttggaa 180
gaagaattga agccattgga agaagttttg aacttggctc aatctaagaa cttccacttg 240
agaccaagag atttgatttc taacattaac gttattgttt tggaattgaa gggttctgaa 300
actactttta tgtgcgagta cgcagacgaa actgctacta tcgttgagtt cttaaataagg 360
tggatcactt tctgccaatc tattattttct actttgacat aa 402
```

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US01/12008

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) : C07K 1/00; A01N 37/18

US CL : 530/350; 514/2

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN: MEDLINE BIOSIS BIOTECHDS EMBASE CAPLUS

WEST

STIC COMMERCIAL DATABASE SEQUENCE SEARCH

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	WO 93/15199 A1 (RHONE-POULENC RORER S.A.) 05 August 1993, see abstract; Fig. 1, page 3, lines 5 and 6, page 4, lines 25-31, and enclosed sequence alignment.	1-4, 6, 8, 18, 20-22, 38 ----- 5, 7, 19
X -- Y	WO 93/15211 A1 (RHONE-POULENC RORER S.A.) 05 August 1993, see abstract, Fig. 1, and enclosed sequence alignment.	1-4, 6, 8, 18, 20-22, 38 ----- 5, 7, 19
Y	WO 96/18412 A1 (BETH ISRAEL HOSPITAL ASSOCIATION) 20 June 1996, See abstract, page 8, lines 15-24, and 31-35; and page 9, lines 21 and 22.	1-8, 18-22, 38



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

09 JULY 2001

Date of mailing of the international search report

02 AUG 2001

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

TERRY J. DEY  
PARALEGAL SPECIALIST  
RICHARD SCHNIZER TECHNOLOGY CENTER 1600

Telephone No. (703) 305-5441

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US01/12008

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	YEH et al. Design of yeast-secreted albumin derivatives for human therapy. Prc. Nat. Acad. Sci. USA. March 1992. Vol. 69, pages 1904-1908, see entire document, especially abstract, page 1905, column 1, lines 14-17 and Fig. 1, panel A, and page 1096, column 1, lines 13-15 of first full paragraph.	1-4, 6, 8, 18, 20-22, 38 ----- 5, 7, 18-22
Y	Database MEDLINE, Accession No. 1999248670, LEE et al. Preparation and characterization of polyethylene-glycol-modified salmon calcitonins. Pharm. Dev. Tech. May 1999. Vol. 4, No. 2, pages 269-275, abstract only.	1-8, 18-22, and 38
Y	Database MEDLINE, Accession No. 97290787, TAKAHASHI et al. Production of bioactive salmon calcitonin from the nonendocrine cell lines COS-7 and CHO. Peptides (1997). Vol. 18, no. 3, pages 439-444, abstract only.	1-8, 18-22, 38

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US01/12008

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 9-17, 23-37, 39-41  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-8, 18-22, 38

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/12008

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

calcitonin, growth hormone releasing factor, IL-2, IL-2 fusion protein, IGF-1, interferon beta, and parathyroid hormone.

The claims are deemed to correspond to the species listed above in the following manner:

Claims 1-6, 8, 18-22, and 38 are generic to all species. Claim 7 is generic to the species IL-2.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: The technical feature which links the species is that they are members of the genus of therapeutic proteins. Claim 1 is drawn broadly to an albumin fusion protein comprising therapeutic protein X. This invention does not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, it lacks a special technical feature because it fails to make a contribution over the prior art. For example, WO 93/15199 discloses therapeutic proteins, such as interleukins and interferons, fused to the albumin of SEQ ID NO:18 of the instant application. Because the invention as a whole fails to make a contribution over the prior art, the technical feature linking the claimed species cannot be a special technical feature under PCT Rule 13.2.



CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number  
**WO 01/79258 A1**

(51) International Patent Classification<sup>7</sup>: C07K 1/00,  
A01N 37/18

(74) Agent: VAN HORN, Charles, E.; Finnegan, Henderson,  
Farabow, Garrett & Dunner, L.L.P., 1300 I Street, N.W.,  
Washington, DC 20005-3315 (US).

(21) International Application Number: PCT/US01/12008

(22) International Filing Date: 12 April 2001 (12.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/229,358 12 April 2000 (12.04.2000) US  
60/199,384 25 April 2000 (25.04.2000) US  
60/256,931 21 December 2000 (21.12.2000) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,  
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,  
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicants (*for all designated States except US*): HU-  
MAN GENOME SCIENCES, INC. [US/US]; 9410 Key  
West Avenue, Rockville, MD 20850 (US). PRINCIPIA  
PHARMACEUTICAL CORPORATION [US/US];  
2650 Eisenhower Avenue, Building C, Norristown, PA  
19403 (US).

**Published:**

- with international search report
- with (an) indication(s) in relation to deposited biological  
material furnished under Rule 13bis separately from the  
description

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): ROSEN, Craig,  
A. [US/US]; 22400 Rolling Hill Lane, Laytonsville,  
MD 20882 (US). SADEGHI, Homayoun [US/US];  
320 E. Court Street, Doylestown, PA 18901 (US).  
PRIOR, Christopher, P. [US/US]; 460 Wyldhaven  
Road, Rosemont, PA 19010 (US). TURNER, Andrew,  
John [GB/US]; Apartment C-28, 305 Conestoga Way,  
Eagleville, PA 19408 (US).

(48) Date of publication of this corrected version:

28 February 2002

(15) Information about Correction:

see PCT Gazette No. 09/2002 of 28 February 2002, Section  
II

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: ALBUMIN FUSION PROTEINS

(57) Abstract: The present invention encompasses albumin fusion proteins. Nucleic acid molecules encoding the albumin fusion proteins of the invention are also encompassed by the invention, as are vectors containing these nucleic acids, host cells transformed with these nucleic acids vectors, and methods of making the albumin fusion proteins of the invention and using these nucleic acids, vectors, and/or host cells. Additionally the present invention encompasses pharmaceutical compositions comprising albumin fusion proteins and methods of treating, preventing, or ameliorating diseases, disorders or conditions using albumin fusion proteins of the invention.



WO 01/79258 A1

## ALBUMIN FUSION PROTEINS

5

10

## BACKGROUND OF THE INVENTION

The invention relates generally to Therapeutic proteins (including, but not limited to, a polypeptide, antibody, or peptide, or fragments and variants thereof) fused to albumin or fragments or variants of albumin. The invention further relates to Therapeutic proteins (including, but not limited to, a polypeptide, antibody, or peptide, or fragments and variants thereof) fused to albumin or fragments or variants of albumin, that exhibit extended shelf-life and/or extended or therapeutic activity in solution. These fusion proteins are herein collectively referred to as "albumin fusion proteins of the invention." The invention encompasses therapeutic albumin fusion proteins, compositions, pharmaceutical compositions, formulations and kits. Nucleic acid molecules encoding the albumin fusion proteins of the invention are also encompassed by the invention, as are vectors containing these nucleic acids, host cells transformed with these nucleic acids vectors, and methods of making the albumin fusion proteins of the invention using these nucleic acids, vectors, and/or host cells.

The invention is also directed to methods of *in vitro* stabilizing a Therapeutic protein via fusion or conjugation of the Therapeutic protein to albumin or fragments or variants of albumin.

Human serum albumin (HSA, or HA), a protein of 585 amino acids in its mature form (as shown in Figure 15 or in SEQ ID NO:18), is responsible for a significant proportion of the osmotic pressure of serum and also functions as a carrier of endogenous and exogenous ligands. At present, HA for clinical use is produced by extraction from human blood. The production of recombinant HA (rHA) in microorganisms has been disclosed in EP 330 451 and EP 361 991.

The role of albumin as a carrier molecule and its inert nature are desirable properties for use as a carrier and transporter of polypeptides *in vivo*. The use of albumin as a component of an albumin fusion protein as a carrier for various proteins has been suggested in WO 93/15199, WO 93/15200, and EP 413 622. The use of N-terminal fragments of HA

for fusions to polypeptides has also been proposed (EP 399 666). Fusion of albumin to the Therapeutic protein may be achieved by genetic manipulation, such that the DNA coding for HA, or a fragment thereof, is joined to the DNA coding for the Therapeutic protein. A suitable host is then transformed or transfected with the fused nucleotide sequences, so arranged on a suitable plasmid as to express a fusion polypeptide. The expression may be effected *in vitro* from, for example, prokaryotic or eukaryotic cells, or *in vivo* e.g. from a transgenic organism.

Therapeutic proteins in their native state or when recombinantly produced, such as interferons and growth hormones, are typically labile molecules exhibiting short shelf-lives, particularly when formulated in aqueous solutions. The instability in these molecules when formulated for administration dictates that many of the molecules must be lyophilized and refrigerated at all times during storage, thereby rendering the molecules difficult to transport and/or store. Storage problems are particularly acute when pharmaceutical formulations must be stored and dispensed outside of the hospital environment. Many protein and peptide drugs also require the addition of high concentrations of other protein such as albumin to reduce or prevent loss of protein due to binding to the container. This is a major concern with respect to proteins such as IFN. For this reason, many Therapeutic proteins are formulated in combination with large proportion of albumin carrier molecule (100-1000 fold excess), though this is an undesirable and expensive feature of the formulation.

Few practical solutions to the storage problems of labile protein molecules have been proposed. Accordingly, there is a need for stabilized, long lasting formulations of proteinaceous therapeutic molecules that are easily dispensed, preferably with a simple formulation requiring minimal post-storage manipulation.

## SUMMARY OF THE INVENTION

The present invention is based, in part, on the discovery that Therapeutic proteins may be stabilized to extend the shelf-life, and/or to retain the Therapeutic protein's activity for extended periods of time in solution, *in vitro* and/or *in vivo*, by genetically or chemically fusing or conjugating the Therapeutic protein to albumin or a fragment (portion) or variant of albumin, that is sufficient to stabilize the protein and/or its activity. In addition it has been determined that the use of albumin-fusion proteins or albumin conjugated proteins may reduce the need to formulate protein solutions with large excesses of carrier proteins (such as albumin, unfused) to prevent loss of Therapeutic proteins due to factors such as binding to the container.

The present invention encompasses albumin fusion proteins comprising a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) fused to albumin or a fragment (portion) or variant of albumin. The present invention also

encompasses albumin fusion proteins comprising a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) fused to albumin or a fragment (portion) or variant of albumin, that is sufficient to prolong the shelf life of the Therapeutic protein, and/or stabilize the Therapeutic protein and/or its activity in solution (or in a pharmaceutical composition) *in vitro* and/or *in vivo*. Nucleic acid molecules encoding the albumin fusion proteins of the invention are also encompassed by the invention, as are vectors containing these nucleic acids, host cells transformed with these nucleic acids vectors, and methods of making the albumin fusion proteins of the invention and using these nucleic acids, vectors, and/or host cells.

The invention also encompasses pharmaceutical formulations comprising an albumin fusion protein of the invention and a pharmaceutically acceptable diluent or carrier. Such formulations may be in a kit or container. Such kit or container may be packaged with instructions pertaining to the extended shelf life of the Therapeutic protein. Such formulations may be used in methods of treating, preventing, ameliotating or diagnosing a disease or disease symptom in a patient, preferably a mammal, most preferably a human, comprising the step of administering the pharmaceutical formulation to the patient.

In other embodiments, the present invention encompasses methods of preventing treating, or ameliorating a disease or disorder. In preferred embodiments, the present invention encompasses a method of treating a disease or disorder listed in the "Preferred Indication Y" column of Table 1 comprising administering to a patient in which such treatment, prevention or amelioration is desired an albumin fusion protein of the invention that comprises a Therapeutic protein portion corresponding to a Therapeutic protein (or fragment or variant thereof) disclosed in the "Therapeutic Protein X" column of Table 1 (in the same row as the disease or disorder to be treated is listed in the "Preferred Indication Y" column of Table 1) in an amount effective to treat prevent or ameliorate the disease or disorder.

In another embodiment, the invention includes a method of extending the shelf life of a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) comprising the step of fusing or conjugating the Therapeutic protein to albumin or a fragment (portion) or variant of albumin, that is sufficient to extend the shelf-life of the Therapeutic protein. In a preferred embodiment, the Therapeutic protein used according to this method is fused to the albumin, or the fragment or variant of albumin. In a most preferred embodiment, the Therapeutic protein used according to this method is fused to albumin, or a fragment or variant of albumin, via recombinant DNA technology or genetic engineering.

In another embodiment, the invention includes a method of stabilizing a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) in solution, comprising the step of fusing or conjugating the Therapeutic protein to albumin or a

fragment (portion) or variant of albumin, that is sufficient to stabilize the Therapeutic protein.

In a preferred embodiment, the Therapeutic protein used according to this method is fused to the albumin, or the fragment or variant of albumin. In a most preferred embodiment, the Therapeutic protein used according to this method is fused to albumin, or a fragment or variant of albumin, via recombinant DNA technology or genetic engineering.

The present invention further includes transgenic organisms modified to contain the nucleic acid molecules of the invention, preferably modified to express the albumin fusion proteins encoded by the nucleic acid molecules.

#### BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts the extended shelf-life of an HA fusion protein in terms of the biological activity (Nb2 cell proliferation) of HA-hGH remaining after incubation in cell culture media for up to 5 weeks at 37°C. Under these conditions, hGH has no observed activity by week 2.

Figure 2 depicts the extended shelf-life of an HA fusion protein in terms of the stable biological activity (Nb2 cell proliferation) of HA-hGH remaining after incubation in cell culture media for up to 3 weeks at 4, 37, or 50°C. Data is normalized to the biological activity of hGH at time zero.

Figures 3A and 3B compare the biological activity of HA-hGH with hGH in the Nb2 cell proliferation assay. Figure 3A shows proliferation after 24 hours of incubation with various concentrations of hGH or the albumin fusion protein, and Figure 3B shows proliferation after 48 hours of incubation with various concentrations of hGH or the albumin fusion protein.

Figure 4 shows a map of a plasmid (pPPC0005) that can be used as the base vector into which polynucleotides encoding the Therapeutic proteins (including polypeptides and fragments and variants thereof) may be cloned to form HA-fusions. Plasmid Map key: PRB1p: *PRB1 S. cerevisiae* promoter; FL: Fusion leader sequence; rHA: cDNA encoding HA; ADH1t: *ADH1 S. cerevisiae* terminator; T3: T3 sequencing primer site; T7: T7 sequencing primer site; Amp R:  $\beta$ -lactamase gene; ori: origin of replication. Please note that in the provisional applications to which this application claims priority, the plasmid in Figure 4 was labeled pPPC0006, instead of pPPC0005. In addition the drawing of this plasmid did not show certain pertinent restriction sites in this vector. Thus in the present application, the drawing is labeled pPPC0005 and more restriction sites of the same vector are shown.

Figure 5 compares the recovery of vial-stored HA-IFN solutions of various concentrations with a stock solution after 48 or 72 hours of storage.

Figure 6 compares the activity of an HA- $\alpha$ -IFN fusion protein after administration to monkeys via IV or SC.

Figure 7 describes the bioavailability and stability of an HA- $\alpha$ -IFN fusion protein.

Figure 8 is a map of an expression vector for the production of HA- $\alpha$ -IFN.

5 Figure 9 shows the location of loops in HA.

Figure 10 is an example of the modification of an HA loop.

Figure 11 is a representation of the HA loops.

Figure 12 shows the HA loop IV.

Figure 13 shows the tertiary structure of HA.

10 Figure 14 shows an example of a scFv-HA fusion

Figure 15 shows the amino acid sequence of the mature form of human albumin (SEQ ID NO:18) and a polynucleotide encoding it (SEQ ID NO:17).

#### DETAILED DESCRIPTION

15 As described above, the present invention is based, in part, on the discovery that a Therapeutic protein (e.g., a polypeptide, antibody, or peptide, or fragments and variants thereof) may be stabilized to extend the shelf-life and/or retain the Therapeutic protein's activity for extended periods of time in solution (or in a pharmaceutical composition) *in vitro* and/or *in vivo*, by genetically fusing or chemically conjugating the Therapeutic protein,  
20 polypeptide or peptide to all or a portion of albumin sufficient to stabilize the protein and its activity.

The present invention relates generally to albumin fusion proteins and methods of treating, preventing, or ameliorating diseases or disorders. As used herein, "albumin fusion protein" refers to a protein formed by the fusion of at least one molecule of albumin (or a  
25 fragment or variant thereof) to at least one molecule of a Therapeutic protein (or fragment or variant thereof). An albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion (i.e., the albumin fusion protein is generated by translation of a nucleic acid in which a polynucleotide encoding all or a  
30 portion of a Therapeutic protein is joined in-frame with a polynucleotide encoding all or a portion of albumin) or chemical conjugation to one another. The Therapeutic protein and albumin protein, once part of the albumin fusion protein, may be referred to as a "portion", "region" or "moiety" of the albumin fusion protein (e.g., a "Therapeutic protein portion" or an "albumin protein portion").

35 In one embodiment, the invention provides an albumin fusion protein comprising, or

alternatively consisting of, a Therapeutic protein (e.g., as described in Table 1) and a serum albumin protein. In other embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment of a Therapeutic protein and a serum albumin protein. In other embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active variant of a Therapeutic protein and a serum albumin protein. In preferred embodiments, the serum albumin protein component of the albumin fusion protein is the mature portion of serum albumin.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein, and a biologically active and/or therapeutically active fragment of serum albumin. In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein and a biologically active and/or therapeutically active variant of serum albumin. In preferred embodiments, the Therapeutic protein portion of the albumin fusion protein is the mature portion of the Therapeutic protein. In a further preferred embodiment, the Therapeutic protein portion of the albumin fusion protein is the extracellular soluble domain of the Therapeutic protein. In an alternative embodiment, the Therapeutic protein portion of the albumin fusion protein is the active form of the Therapeutic protein.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment or variant of a Therapeutic protein and a biologically active and/or therapeutically active fragment or variant of serum albumin. In preferred embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, the mature portion of a Therapeutic protein and the mature portion of serum albumin.

### **Therapeutic proteins**

As stated above, an albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion or chemical conjugation.

As used herein, "Therapeutic protein" refers to proteins, polypeptides, antibodies, peptides or fragments or variants thereof, having one or more therapeutic and/or biological activities. Therapeutic proteins encompassed by the invention include but are not limited to, proteins, polypeptides, peptides, antibodies, and biologics. (The terms peptides, proteins, and polypeptides are used interchangeably herein.) It is specifically contemplated that the term "Therapeutic protein" encompasses antibodies and fragments and variants thereof. Thus an albumin fusion protein of the invention may contain at least a fragment or variant of a

Therapeutic protein, and/or at least a fragment or variant of an antibody. Additionally, the term "Therapeutic protein" may refer to the endogenous or naturally occurring correlate of a Therapeutic protein.

By a polypeptide displaying a "therapeutic activity" or a protein that is "therapeutically active" is meant a polypeptide that possesses one or more known biological and/or therapeutic activities associated with a Therapeutic protein such as one or more of the Therapeutic proteins described herein or otherwise known in the art. As a non-limiting example, a "Therapeutic protein" is a protein that is useful to treat, prevent or ameliorate a disease, condition or disorder. As a non-limiting example, a "Therapeutic protein" may be one that binds specifically to a particular cell type (normal (e.g., lymphocytes) or abnormal e.g., (cancer cells)) and therefore may be used to target a compound (drug, or cytotoxic agent) to that cell type specifically.

In another non-limiting example, a "Therapeutic protein" is a protein that has a biological activity, and in particular, a biological activity that is useful for treating preventing or ameliorating a disease. A non-inclusive list of biological activities that may be possessed by a Therapeutic protein includes, enhancing the immune response, promoting angiogenesis, inhibiting angiogenesis, regulating hematopoietic functions, stimulating nerve growth, enhancing an immune response, inhibiting an immune response, or any one or more of the biological activities described in the "Biological Activities" section below.

As used herein, "therapeutic activity" or "activity" may refer to an activity whose effect is consistent with a desirable therapeutic outcome in humans, or to desired effects in non-human mammals or in other species or organisms. Therapeutic activity may be measured *in vivo* or *in vitro*. For example, a desirable effect may be assayed in cell culture. As an example, when hGH is the Therapeutic protein, the effects of hGH on cell proliferation as described in Example 1 may be used as the endpoint for which therapeutic activity is measured. Such *in vitro* or cell culture assays are commonly available for many Therapeutic proteins as described in the art.

Examples of useful assays for particular Therapeutic proteins include, but are not limited to: IFN $\beta$  (anti-viral assay: Rubinstein et al., J. Virol. 37(2):755-8 (1981); Anti-proliferation assay: Gao Y, et al., Mol Cell Biol. 19(11):7305-13 (1999; and bioassay: Czarniecki et al., J. Virol. 49 p490 (1984)); IGF-1 (bioassay: Karey et al., Cancer Research 48 p4083); IL-2 soluble receptor and IL-2 receptor antagonist (bioassay: Symons et al., Lymphokines and Interferons, A practical Approach, Clemens et al., eds. IRL Press, p272; T cell proliferation assay: Science 223: 1412-1415 (1984) natural killer (NK) cell and CTL cytotoxicity assay: Science 288: 675-678 (2000); CTLL-2 Proliferation: Gillis et al J. Immunol. 120:2027 (1978)); calcitonin (assays include, but are not limited to, Hypocalcemic



Rat Bioassay, bone resorbing assay and the pit assay, CT receptor binding assay, CAMP stimulation assay: J Bone Miner Res 14(8):1425-31 (1999)); growth hormone releasing factor (Biological assay: Mol Endocrinol 6(10):1734-44, Mol Endocrinol, 7:77-84 (1992)); parathyroid hormone (adenyl cyclase stimulation in rat osteosarcoma cells, ovariectomized rat model of osteoporosis: IUBMB Life 49(2):131-5 (2000)), or as shown in Table 1 in the "Exemplary Activity Assay" column.

Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, such as cell surface and secretory proteins, are often modified by the attachment of one or more oligosaccharide groups. The modification, referred to as glycosylation, can dramatically affect the physical properties of proteins and can be important in protein stability, secretion, and localization. Glycosylation occurs at specific locations along the polypeptide backbone. There are usually two major types of glycosylation: glycosylation characterized by O-linked oligosaccharides, which are attached to serine or threonine residues; and glycosylation characterized by N-linked oligosaccharides, which are attached to asparagine residues in an Asn-X-Ser/Thr sequence, where X can be any amino acid except proline. N-acetylneuramic acid (also known as sialic acid) is usually the terminal residue of both N-linked and O-linked oligosaccharides. Variables such as protein structure and cell type influence the number and nature of the carbohydrate units within the chains at different glycosylation sites. Glycosylation isomers are also common at the same site within a given cell type.

For example, several types of human interferon are glycosylated. Natural human interferon- $\alpha 2$  is O-glycosylated at threonine 106, and N-glycosylation occurs at asparagine 72 in interferon- $\alpha 14$  (Adolf *et al.*, J. Biochem 276:511 (1991); Nyman TA *et al.*, J. Biochem 329:295 (1998)). The oligosaccharides at asparagine 80 in natural interferon- $\beta 1\alpha$  may play an important factor in the solubility and stability of the protein, but may not be essential for its biological activity. This permits the production of an unglycosylated analog (interferon- $\beta 1b$ ) engineered with sequence modifications to enhance stability (Hosoi *et al.*, J. Interferon Res. 8:375 (1988; Karpusas *et al.*, Cell Mol Life Sci 54:1203 (1998); Knight, J. Interferon Res. 2:421 (1982); Runkel *et al.*, Pharm Res 15:641 (1998); Lin, Dev. Biol. Stand. 96:97 (1998))1. Interferon- $\gamma$  contains two N-linked oligosaccharide chains at positions 25 and 97, both important for the efficient formation of the bioactive recombinant protein, and having an influence on the pharmacokinetic properties of the protein (Sareneva *et al.*, Eur. J. Biochem 242:191 (1996); Sareneva *et al.*, Biochem J. 303:831 (1994); Sareneva *et al.*, J. Interferon

Res. 13:267 (1993)). Mixed O-linked and N-linked glycosylation also occurs, for example in human erythropoietin, N-linked glycosylation occurs at asparagine residues located at positions 24, 38 and 83 while O-linked glycosylation occurs at a serine residue located at position 126 (Lai *et al.*, J. Biol. Chem. 261:3116 (1986); Broudy *et al.*, Arch. Biochem. Biophys. 265:329 (1988)).

Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, as well as analogs and variants thereof, may be modified so that glycosylation at one or more sites is altered as a result of manipulation(s) of their nucleic acid sequence, by the host cell in which they are expressed, or due to other conditions of their expression. For example, glycosylation isomers may be produced by abolishing or introducing glycosylation sites, *e.g.*, by substitution or deletion of amino acid residues, such as substitution of glutamine for asparagine, or unglycosylated recombinant proteins may be produced by expressing the proteins in host cells that will not glycosylate them, *e.g.* in *E. coli* or glycosylation-deficient yeast. These approaches are described in more detail below and are known in the art.

Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention include, but are not limited to, calcitonin, growth hormone releasing factor, insulin-like growth factor-1, interferon beta, interleukin-2, and parathyroid hormone, in addition to those described in Table 1. These proteins and nucleic acid sequences encoding these proteins are well known and available in public databases such as Chemical Abstracts Services Databases (*e.g.*, the CAS Registry), GenBank, and GenSeq as shown in Table 1.

Additional Therapeutic proteins corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention include, but are not limited to, one or more of the Therapeutic proteins or peptides disclosed in the "Therapeutic Protein X" column of Table 1, or fragment or variable thereof.

Table 1 provides a non-exhaustive list of Therapeutic proteins that correspond to a Therapeutic protein portion of an albumin fusion protein of the invention. The "Therapeutic Protein X" column discloses Therapeutic protein molecules followed by parentheses containing scientific and brand names that comprise, or alternatively consist of, that Therapeutic protein molecule or a fragment or variant thereof. "Therapeutic protein X" as used herein may refer either to an individual Therapeutic protein molecule (as defined by the amino acid sequence obtainable from the CAS and Genbank accession numbers), or to the entire group of Therapeutic proteins associated with a given Therapeutic protein molecule disclosed in this column. The "Exemplary Identifier" column provides Chemical Abstracts Services (CAS) Registry Numbers (published by the American Chemical Society) and/or Genbank Accession Numbers ((*e.g.*, Locus ID, NP\_XXXXX (Reference Sequence Protein),

and XP\_XXXXXX (Model Protein) identifiers available through the national Center for Biotechnology Information (NCBI) webpage at [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) that correspond to entries in the CAS Registry or Genbank database which contain an amino acid sequence of the Therapeutic Protein Molecule or of a fragment or variant of the Therapeutic Protein Molecule.

5 The summary pages associated with each of these CAS and Genbank Accession Numbers are each incorporated by reference in their entireties, particularly with respect to the amino acid sequences described therein. The "PCT/Patent Reference" column provides U.S. Patent numbers, or PCT International Publication Numbers corresponding to patents and/or published patent applications that describe the Therapeutic protein molecule. Each of the  
10 patents and/or published patent applications cited in the "PCT/Patent Reference" column are herein incorporated by reference in their entireties. In particular, the amino acid sequences of the specified polypeptide set forth in the sequence listing of each cited "PCT/Patent Reference", the variants of these amino acid sequences (mutations, fragments, etc.) set forth, for example, in the detailed description of each cited "PCT/Patent Reference", the therapeutic  
15 indications set forth, for example, in the detailed description of each cited "PCT/Patent Reference", and the activity assays for the specified polypeptide set forth in the detailed description, and more particularly, the examples of each cited "PCT/Patent Reference" are incorporated herein by reference. The "Biological activity" column describes Biological activities associated with the Therapeutic protein molecule. The "Exemplary Activity Assay"  
20 column provides references that describe assays which may be used to test the therapeutic and/or biological activity of a Therapeutic protein or an albumin fusion protein of the invention comprising a Therapeutic protein X portion. Each of the references cited in the "Exemplary Activity Assay" column are herein incorporated by reference in their entireties, particularly with respect to the description of the respective activity assay described in the  
25 reference (see Methods section, for example) for assaying the corresponding biological activity set forth in the "Biological Activity" column of Table 1. The "Preferred Indication Y" column describes disease, disorders, and/or conditions that may be treated, prevented, diagnosed, or ameliorated by Therapeutic protein X or an albumin fusion protein of the invention comprising a Therapeutic protein X portion.

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Calcitonin (Salmon Calcitonin (Salcatonin)); Calcitonin human-salmon hybrid; Forcaltonin; Fortical; Calcitonin; Calcitonina Almirall; Calcitonina Hubber; Calcimar; Calsynar; Calogen; Miacalcic; Miacalcin; SB205614; Macritonin; Cibacalcin; Cibacalcina; Cibacalcine; Salmocalcin; Powderject (Calcitonin)	CAS-21215-62-3 CAS-47931-85-1 CAS-152874-80-1 CAS-60731-46-6 LocusID:796 NP_001732 XP_006209	EP70675-A JP05255391-A EP95351-A EP95351-A WO9415962-A US5858978-A	Calcitonin is a peptide hormone synthesized by the parafollicular cells of the thyroid. It causes reduction in serum calcium--an effect opposite to that of human parathyroid hormone.	Hypocalcemic Rat Bioassay, bone resorbing assay and the pit assay, CT receptor binding assay, CAMP stimulation assay; J Bone Miner Res 1999 Aug;14(8):1425-31	Bone Disorders; Fracture prevention; Hypercalcemia; Malignant hypercalcemia; Osteoporosis; Paget's disease
Growth hormone releasing factor (Sermorelin acetate; Pralmorelin; Somatorelin; Somatotoliberin; Geref; Groliberin)	CAS-114466-38-5 CAS-86168-78-7 CAS-158827-34-0 CAS-83930-13-6 CAS-9034-39-3 LocusID:2691 NP_066567 XP_012959	WO8605812 AU8429733 US5700775 US5939386 US5416073 US4888286	Growth hormone-releasing factor (GRF) is released by the hypothalamus and acts on the adenohypophyse to stimulate the production of growth hormone.	Growth hormone-releasing peptides (GHRPs) are known to release growth hormone (GH) in vivo and in vitro by a direct action on receptors in anterior pituitary cells. Biological activity can be measured in cell lines expressing growth hormone releasing factor receptor. (Mol Endocrinol 1992 Oct;6(10):1734-44, Molecular Endocrinology, Vol 7, 77-84).	Cachexia; Diagnostic Agents; Female Infertility; Growth hormone deficiency; Muscle Wasting; Short Stature; Postmenopausal Osteoporosis; Idiopathic growth hormone deficiency

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
IL-2 (Aldesleukin; interleukin-2 fusion toxin; T cell growth factor; PROLEUKIN; IMMUNACE; CELEUK; ONCOLIPIN 2; MACROLIN)	CAS-110942-02-4 LocusID:3558 NP_000577 XP_003373	EP91539-A EP215576-A WO9834952-A2 WO9926663-A2	IL-2 is a powerfully immunoregulatory lymphokine that is produced by lectin- or antigen-activated T cells that promotes the growth of B and T cells and augments NK cell and CTL cell killing activity.	T cell proliferation assay "Biological activity of recombinant human interleukin-2 produced in Escherichia coli." Science 223: 1412-1415, 1984. natural killer (NK) cell and CTL cytotoxicity assay "Control of homeostasis of CD8+ memory T cells by opposing cytokines. Science 288: 675-678, 2000; CTLL-2 Proliferation : Gillis et al (1978) J. Immunol. 120, 2027	Metastatic renal cell carcinoma and metastatic melanoma; Malignant Melanoma; Renal cell carcinoma; HIV infections treatment; Inflammatory bowel disorders; Kaposi's sarcoma; Leukaemia; Multiple sclerosis; Rheumatoid arthritis; Transplant rejection; Type I diabetes mellitus; Lung Cancer; Acute myeloid leukaemia; Hepatitis C; Non-Hodgkin's lymphoma; Ovarian cancer
IL-2 fusion protein (Denileukin diftitox; DAB389interleukin-2; LY 333348; interleukin-2 fusion protein; interleukin-2-fusion toxin; ONTAK)	CAS-173146-27-5 LocusID:3558 NP_000577 XP_003373	EP91539-A EP215576-A WO9834952-A2 WO9926663-A2	IL-2, a cytokine first identified as a T cell growth factor, has been proven to activate many cell types including T cells, B cells, and polymorphonuclear neutrophils (PMN3). Fusion proteins of IL-2 with antibodies with specificity for tumor associated antigens were used to treat cancers by concentrating the cytokine in the tumor microenvironment and in so doing directly enhance the tumoricidal effect of the antibody and/or enhance the host immune response (T-cell, B-cell or NK) against the tumor.	CTLL-2 T cell line proliferation assay for IL-2 bioactivity. (PMID: 11101150); CTLL-2 Proliferation : Gillis et al (1978) J. Immunol. 120, 2027	T-cell lymphoma; Alopecia; Atopic dermatitis; Chronic lymphocytic leukaemia; Cutaneous T-cell lymphoma; HIV infections treatment; Inflammatory bowel disorders; Multiple sclerosis; Non-Hodgkin's lymphoma; Psoriasis; Rheumatoid arthritis; Transplant rejection; Type I diabetes mellitus

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Insulin-like growth factor-1 (Mecasermin; Somazon; IGF-1; IGF-1 complex; CEP 151; CGP 35126; FK 780; Mecar; RHIGF-1; Somatomedin-1; Somatomedin-C; SOMATOKINE; MYOTROPHIN; IGEF; DepoIGF-1)	CAS-68562-41-4; LocusID:3479; NP_000609; XP_012220	WO9516703-A1 EP229750-A WO8600619-A EP155655-A WO9204363-A	IGF-1 is a pleiotropic polypeptide with a wide range of actions in both central and peripheral nervous systems. It is involved in growth and development and protects neurons against cell death via the activation of intracellular pathways implicating phosphatidylinositol 3/Akt kinase.	IGF-I activity may be assayed in vitro using a serum withdrawal apoptosis-protection assay. (J Endocrinol 2000 Oct; 167(1):165-74). Proliferation assay using breast carcinoma cell line MCF-7 (Karey 1988 Cancer Res. 48: 4083)	Diabetes mellitus; Growth disorders; Frailty; Amyotrophic lateral sclerosis; Osteoarthritis; Kidney disease & neuropathy; Dwarfism; HIV-1 infections; Myocardial ischaemia; Osteoporosis; Multiple sclerosis; Nerve disorders; Burns; diabetes; peripheral neuropathies
Interferon beta (Interferon beta-1a; Interferon beta-1b; Interferon-beta-serine; SH 579; ZK 157046; BCDF; beta-2 IF; Interferon-beta-2; rhIL-6; SJ0031; DL 8234; FERON; IFNbeta; BETASERON; AVONEX; REBIF; BETAIFERON; SIGOSIX)	CAS-74899-71-1 CAS-145155-23-3 CAS-145258-61-3 LocusID:3456 LocusID:3457 NP_002167 XP_005410	EP28033-A EP48970-A US5326859-A WO8202715-A US4499188-A WO8302461-A EP237019-A EP146413-A US4518584-A US4588585-A	IFN beta displays anti-viral, anti-proliferative and immunomodulatory activities including modulation of MHC antigen expression, NK cell activity and IFN $\gamma$ production and IL12 production in monocytes.	Anti-viral assay: Rubinstein S, Familletti PC, Pestka S. (1981) Convenient assay for interferons. J. Virol. 37(2):755-8 - Anti-proliferation assay: Gao Y, et al (1999) Sensitivity of an Epstein-Barr virus-positive tumor line, Daudi, to alpha interferon correlates with expression of a GC-rich viral transcript. Mol Cell Biol. 19(11):7305-13.	Melanoma; solid tumors; Multiple sclerosis; bacterial infections; chemoprotection; thrombocytopenia; HIV infections; prostate cancer; cancer; hematological malignancies; hematological disorders; preleukemia; glioma; hepatitis B; hepatitis C; human papillomavirus; pulmonary fibrosis; Age-related macular degeneration; Brain cancer; Glioblastoma multiforme; Liver cancer; Malignant melanoma; Colorectal cancer; Crohn's disease; Neurological disorders; Non-small cell lung cancer; Rheumatoid arthritis; Ulcerative colitis

Therapeutic Protein X	Exemplary Identifier	PCT/Patent Reference	Biological Activity	Exemplary Activity Assay	Preferred Indication Y
Parathyroid hormone (PTH; Ostarabolin; ALX1-11; hPTH 1-34; LY 333334; MN 10T; parathyroid hormone (1-31); FORTEO; PARATHAR)	CAS-9002-64-6 LocusID:5741 NP_000306 XP_005988	DE3312928-A WO8809376-A WO880316-A US5420242-A	Main function is to regulate concentrations of Calcium and phosphate in the blood o Imbalance of Ca/Pi causes altered neural transmission, destruction of bone tissue, hampered bone growth and muscle tetany. Interacts with calcitonin. o stimulates the activity of osteocytes which break down bone tissue and release Ca into blood o enhances absorption of Ca/Pi from small intestine into blood o promotes reabsorption of Ca and inhibits Pi by kidney tubules.	Adenylyl cyclase stimulation in rat osteosarcoma cells, ovariectomized rat model of osteoporosis: IUBMB Life 2000 Feb;49(2):131-5	Osteoporosis; Malignant hypercalcaemia

In preferred embodiments, the albumin fusion proteins of the invention are capable of a therapeutic activity and/or biologic activity corresponding to the therapeutic activity and/or biologic activity of the Therapeutic protein corresponding to the Therapeutic protein portion of the albumin fusion protein listed in the corresponding row of Table 1. (See, e.g., the "Biological Activity" and "Therapeutic Protein X" columns of Table 1.) In further preferred embodiments, the therapeutically active protein portions of the albumin fusion proteins of the invention are fragments or variants of the reference sequence cited in the "Exemplary Identifier" column of Table 1, and are capable of the therapeutic activity and/or biologic activity of the corresponding Therapeutic protein disclosed in "Biological Activity" column of Table 1.

### ***Polypeptide and Polynucleotide Fragments and Variants***

#### ***Fragments***

The present invention is further directed to fragments of the Therapeutic proteins described in Table 1, albumin proteins, and/or albumin fusion proteins of the invention.

Even if deletion of one or more amino acids from the N-terminus of a protein results in modification or loss of one or more biological functions of the Therapeutic protein, albumin protein, and/or albumin fusion protein, other Therapeutic activities and/or functional activities (e.g., biological activities, ability to multimerize, ability to bind a ligand) may still be retained.

For example, the ability of polypeptides with N-terminal deletions to induce and/or bind to antibodies which recognize the complete or mature forms of the polypeptides generally will be retained when less than the majority of the residues of the complete polypeptide are removed from the N-terminus. Whether a particular polypeptide lacking N-terminal residues of a complete polypeptide retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art. It is not unlikely that a mutein with a large number of deleted N-terminal amino acid residues may retain some biological or immunogenic activities. In fact, peptides composed of as few as six amino acid residues may often evoke an immune response.

Accordingly, fragments of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, include the full length protein as well as polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the reference polypeptide (e.g., a Therapeutic protein as disclosed in Table 1). In particular, N-terminal deletions may be described by the general formula  $m-q$ , where  $q$  is a whole integer representing the total number of amino acid residues in a reference polypeptide (e.g., a Therapeutic protein referred to in Table 1), and  $m$  is defined as any



integer ranging from 2 to  $q-6$ . Polynucleotides encoding these polypeptides are also encompassed by the invention.

In addition, fragments of serum albumin polypeptides corresponding to an albumin protein portion of an albumin fusion protein of the invention, include the full length protein as well as polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the reference polypeptide (i.e., serum albumin). In particular, N-terminal deletions may be described by the general formula  $m-585$ , where 585 is a whole integer representing the total number of amino acid residues in serum albumin (SEQ ID NO:18), and  $m$  is defined as any integer ranging from 2 to 579. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Moreover, fragments of albumin fusion proteins of the invention, include the full length albumin fusion protein as well as polypeptides having one or more residues deleted from the amino terminus of the albumin fusion protein. In particular, N-terminal deletions may be described by the general formula  $m-q$ , where  $q$  is a whole integer representing the total number of amino acid residues in the albumin fusion protein, and  $m$  is defined as any integer ranging from 2 to  $q-6$ . Polynucleotides encoding these polypeptides are also encompassed by the invention.

Also as mentioned above, even if deletion of one or more amino acids from the N-terminus or C-terminus of a reference polypeptide (e.g., a Therapeutic protein and/or serum albumin protein) results in modification or loss of one or more biological functions of the protein, other functional activities (e.g., biological activities, ability to multimerize, ability to bind a ligand) and/or Therapeutic activities may still be retained. For example the ability of polypeptides with C-terminal deletions to induce and/or bind to antibodies which recognize the complete or mature forms of the polypeptide generally will be retained when less than the majority of the residues of the complete or mature polypeptide are removed from the C-terminus. Whether a particular polypeptide lacking the N-terminal and/or C-terminal residues of a reference polypeptide retains Therapeutic activity can readily be determined by routine methods described herein and/or otherwise known in the art.

The present invention further provides polypeptides having one or more residues deleted from the carboxy terminus of the amino acid sequence of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention (e.g., a Therapeutic protein referred to in Table 1). In particular, C-terminal deletions may be described by the general formula  $1-n$ , where  $n$  is any whole integer ranging from 6 to  $q-1$ , and where  $q$  is a whole integer representing the total number of amino acid residues in a reference polypeptide (e.g., a Therapeutic protein referred to in Table 1). Polynucleotides encoding these polypeptides are also encompassed by the invention.

In addition, the present invention provides polypeptides having one or more residues deleted from the carboxy terminus of the amino acid sequence of an albumin protein corresponding to an albumin protein portion of an albumin fusion protein of the invention (e.g., serum albumin). In particular, C-terminal deletions may be described by the general formula 1-n, where n is any whole integer ranging from 6 to 584, where 584 is the whole integer representing the total number of amino acid residues in serum albumin (SEQ ID NO:18) minus 1. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Moreover, the present invention provides polypeptides having one or more residues deleted from the carboxy terminus of an albumin fusion protein of the invention. In particular, C-terminal deletions may be described by the general formula 1-n, where n is any whole integer ranging from 6 to q-1, and where q is a whole integer representing the total number of amino acid residues in an albumin fusion protein of the invention. Polynucleotides encoding these polypeptides are also encompassed by the invention.

In addition, any of the above described N- or C-terminal deletions can be combined to produce a N- and C-terminal deleted reference polypeptide. The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues m-n of a reference polypeptide (e.g., a Therapeutic protein referred to in Table 1, or serum albumin (e.g., SEQ ID NO:18), or an albumin fusion protein of the invention) where n and m are integers as described above. Polynucleotides encoding these polypeptides are also encompassed by the invention.

The present application is also directed to proteins containing polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to a reference polypeptide sequence (e.g., a Therapeutic protein, serum albumin protein or an albumin fusion protein of the invention) set forth herein, or fragments thereof. In preferred embodiments, the application is directed to proteins comprising polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to reference polypeptides having the amino acid sequence of N- and C-terminal deletions as described above. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Preferred polypeptide fragments of the invention are fragments comprising, or alternatively, consisting of, an amino acid sequence that displays a Therapeutic activity and/or functional activity (e.g. biological activity) of the polypeptide sequence of the Therapeutic protein or serum albumin protein of which the amino acid sequence is a fragment.

Other preferred polypeptide fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include

an improved desired activity, or a decreased undesirable activity.

#### *Variants*

“Variant” refers to a polynucleotide or nucleic acid differing from a reference nucleic acid or polypeptide, but retaining essential properties thereof. Generally, variants are overall  
5 closely similar, and, in many regions, identical to the reference nucleic acid or polypeptide.

As used herein, “variant”, refers to a Therapeutic protein portion of an albumin fusion protein of the invention, albumin portion of an albumin fusion protein of the invention, or albumin fusion protein differing in sequence from a Therapeutic protein (e.g. see  
10 “therapeutic” column of Table 1), albumin protein, and/or albumin fusion protein of the invention, respectively, but retaining at least one functional and/or therapeutic property thereof (e.g., a therapeutic activity and/or biological activity as disclosed in the “Biological Activity” column of Table 1) as described elsewhere herein or otherwise known in the art. Generally, variants are overall very similar, and, in many regions, identical to the amino acid  
15 sequence of the Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, albumin protein corresponding to an albumin protein portion of an albumin fusion protein of the invention, and/or albumin fusion protein of the invention. Nucleic acids encoding these variants are also encompassed by the invention.

The present invention is also directed to proteins which comprise, or alternatively  
20 consist of, an amino acid sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100%, identical to, for example, the amino acid sequence of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention (e.g., an amino acid sequence disclosed in the “Exemplary Identifier” column of Table 1, or fragments or variants thereof), albumin proteins (e.g., SEQ ID NO:18 or  
25 fragments or variants thereof) corresponding to an albumin protein portion of an albumin fusion protein of the invention, and/or albumin fusion proteins of the invention. Fragments of these polypeptides are also provided (e.g., those fragments described herein). Further polypeptides encompassed by the invention are polypeptides encoded by polynucleotides which hybridize to the complement of a nucleic acid molecule encoding an amino acid  
30 sequence of the invention under stringent hybridization conditions (e.g., hybridization to filter bound DNA in 6X Sodium chloride/Sodium citrate (SSC) at about 45 degrees Celsius, followed by one or more washes in 0.2X SSC, 0.1% SDS at about 50 - 65 degrees Celsius), under highly stringent conditions (e.g., hybridization to filter bound DNA in 6X sodium chloride/Sodium citrate (SSC) at about 45 degrees Celsius, followed by one or more washes  
35 in 0.1X SSC, 0.2% SDS at about 68 degrees Celsius), or under other stringent hybridization conditions which are known to those of skill in the art (see, for example, Ausubel, F.M. et

al., eds., 1989 *Current protocol in Molecular Biology*, Green publishing associates, Inc., and John Wiley & Sons Inc., New York, at pages 6.3.1 - 6.3.6 and 2.10.3). Polynucleotides encoding these polypeptides are also encompassed by the invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, or substituted with another amino acid. These alterations of the reference sequence may occur at the amino- or carboxy-terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence of an albumin fusion protein of the invention or a fragment thereof (such as the Therapeutic protein portion of the albumin fusion protein or the albumin portion of the albumin fusion protein), can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci.6:237-245 (1990)). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is expressed as percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject

residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C- terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The variant will usually have at least 75 % (preferably at least about 80%, 90%, 95% or 99%) sequence identity with a length of normal HA or Therapeutic protein which is the same length as the variant. Homology or identity at the nucleotide or amino acid sequence level is determined by BLAST (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs blastp, blastn, blastx, tblastn and tblastx (Karlin *et al.*, Proc. Natl. Acad. Sci. USA 87: 2264-2268 (1990) and Altschul, J. Mol. Evol. 36: 290-300 (1993), fully incorporated by reference) which are tailored for sequence similarity searching.

The approach used by the BLAST program is to first consider similar segments between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul *et al.*, (Nature Genetics 6: 119-129 (1994)) which is fully

incorporated by reference. The search parameters for histogram, descriptions, alignments, expect (i.e., the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the BLOSUM62 matrix (Henikoff *et al.*, Proc. Natl. Acad. Sci. USA 89: 10915-10919 (1992), fully incorporated by reference). For blastn, the scoring matrix is set by the ratios of M (i.e., the reward score for a pair of matching residues) to N (i.e., the penalty score for mismatching residues), wherein the default values for M and N are 5 and -4, respectively. Four blastn parameters may be adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every wink<sup>th</sup> position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent Blastp parameter settings were Q=9; R=2; wink=1; and gapw=32. A Bestfit comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

The polynucleotide variants of the invention may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, polypeptide variants in which less than 50, less than 40, less than 30, less than 20, less than 10, or 5-50, 5-25, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host, such as, yeast or *E. coli*).

In a preferred embodiment, a polynucleotide encoding an albumin portion of an albumin fusion protein of the invention is optimized for expression in yeast or mammalian cells. In further preferred embodiment, a polynucleotide encoding a Therapeutic protein portion of an albumin fusion protein of the invention is optimized for expression in yeast or mammalian cells. In a still further preferred embodiment, a polynucleotide encoding an albumin fusion protein of the invention is optimized for expression in yeast or mammalian cells.

In an alternative embodiment, a codon optimized polynucleotide encoding a Therapeutic protein portion of an albumin fusion protein of the invention does not hybridize to the wild type polynucleotide encoding the Therapeutic protein under stringent hybridization conditions as described herein. In a further embodiment, a codon optimized polynucleotide encoding an albumin portion of an albumin fusion protein of the invention does not hybridize

to the wild type polynucleotide encoding the albumin protein under stringent hybridization conditions as described herein. In another embodiment, a codon optimized polynucleotide encoding an albumin fusion protein of the invention does not hybridize to the wild type polynucleotide encoding the Therapeutic protein portion or the albumin protein portion under stringent hybridization conditions as described herein.

In an additional embodiment, polynucleotides encoding a Therapeutic protein portion of an albumin fusion protein of the invention do not comprise, or alternatively consist of, the naturally occurring sequence of that Therapeutic protein. In a further embodiment, polynucleotides encoding an albumin protein portion of an albumin fusion protein of the invention do not comprise, or alternatively consist of, the naturally occurring sequence of albumin protein. In an alternative embodiment, polynucleotides encoding an albumin fusion protein of the invention do not comprise, or alternatively consist of, the naturally occurring sequence of a Therapeutic protein portion or the albumin protein portion.

In a preferred embodiment, an IL-2 Therapeutic protein portion of an albumin fusion protein of the invention is encoded by a polynucleotide that has been codon optimized for expression in yeast and having the nucleotide sequence of SEQ ID NO:36. In another preferred embodiment, an IL-2 Therapeutic protein portion of an albumin fusion protein of the invention is encoded by a polynucleotide that has been codon optimized for expression in yeast and having the nucleotide sequence of a portion of the nucleotide sequence of SEQ ID NO:36. In another preferred embodiment, an IL-2 Therapeutic protein portion of an albumin fusion protein of the invention is encoded by a polynucleotide that has been codon optimized for expression in yeast and having the nucleotide sequence of at least 30 contiguous nucleotides of the nucleotide sequence of SEQ ID NO:36. In another preferred embodiment, an IL-2 Therapeutic protein portion of an albumin fusion protein of the invention is encoded by a polynucleotide that has been codon optimized for expression in yeast and having the nucleotide sequence of at least 90 contiguous nucleotides of the nucleotide sequence of SEQ ID NO:36.

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985)). These allelic variants can vary at either the polynucleotide and/or polypeptide level and are included in the present invention. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus

or C-terminus of the polypeptide of the present invention without substantial loss of biological function. As an example, Ron et al. (J. Biol. Chem. 268: 2984-2988 (1993)) reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem. 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which have a functional activity (e.g., biological activity and/or therapeutic activity). In highly preferred embodiments the invention provides variants of albumin fusion proteins that have a functional activity (e.g., biological activity and/or therapeutic activity, such as that disclosed in the "Biological Activity" column in Table 1) that corresponds to one or more biological and/or therapeutic activities of the Therapeutic protein corresponding to the Therapeutic protein portion of the albumin fusion protein. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as to have little effect on activity.

In preferred embodiments, the variants of the invention have conservative substitutions. By "conservative substitutions" is intended swaps within groups such as replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu;



replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

5 Guidance concerning how to make phenotypically silent amino acid substitutions is provided, for example, in Bowie et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science* 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

10 The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function.

Thus, positions tolerating amino acid substitution could be modified while still maintaining 15 biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. See Cunningham and Wells, 20 *Science* 244:1081-1085 (1989). The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most 25 buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic 30 residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly. Besides conservative amino acid substitution, variants of the present invention include (i) polypeptides containing substitutions of one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by 35 the genetic code, or (ii) polypeptides containing substitutions of one or more of the amino acid residues having a substituent group, or (iii) polypeptides which have been fused with or

chemically conjugated to another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), (iv) polypeptide containing additional amino acids, such as, for example, an IgG Fc fusion region peptide, . Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. See Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).

In specific embodiments, the polypeptides of the invention comprise, or alternatively, consist of, fragments or variants of the amino acid sequence of a Therapeutic protein described herein and/or human serum albumin, and/or albumin fusion protein of the invention, wherein the fragments or variants have 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, amino acid residue additions, substitutions, and/or deletions when compared to the reference amino acid sequence. In preferred embodiments, the amino acid substitutions are conservative. Nucleic acids encoding these polypeptides are also encompassed by the invention.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as post-translational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation,

demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-  
5 RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POST-TRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth. Enzymol. 182:626-646  
10 (1990); Rattan et al., Ann. N.Y. Acad. Sci. 663:48-62 (1992)).

#### *Functional activity*

“A polypeptide having functional activity” refers to a polypeptide capable of displaying one or more known functional activities associated with the full-length, pro-  
15 protein, and/or mature form of a Therapeutic protein. Such functional activities include, but are not limited to, biological activity, antigenicity [ability to bind (or compete with a polypeptide for binding) to an anti-polypeptide antibody], immunogenicity (ability to generate antibody which binds to a specific polypeptide of the invention), ability to form multimers with polypeptides of the invention, and ability to bind to a receptor or ligand for a  
20 polypeptide.

“A polypeptide having biological activity” refers to a polypeptide exhibiting activity similar to, but not necessarily identical to, an activity of a Therapeutic protein of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be  
25 identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention).

30 In preferred embodiments, an albumin fusion protein of the invention has at least one biological and/or therapeutic activity associated with the Therapeutic protein (or fragment or variant thereof) when it is not fused to albumin.

The albumin fusion proteins of the invention can be assayed for functional activity (e.g., biological activity) using or routinely modifying assays known in the art, as well as  
35 assays described herein. Specifically, albumin fusion proteins may be assayed for functional activity (e.g., biological activity or therapeutic activity) using the assay referenced in the

“Exemplary Activity Assay” column of Table 1. Additionally, one of skill in the art may routinely assay fragments of a Therapeutic protein corresponding to a Therapeutic protein portion of an albumin fusion protein of the invention, for activity using assays referenced in its corresponding row of Table 1. Further, one of skill in the art may routinely assay  
5 fragments of an albumin protein corresponding to an albumin protein portion of an albumin fusion protein of the invention, for activity using assays known in the art and/or as described in the Examples section below.

For example, in one embodiment where one is assaying for the ability of an albumin fusion protein of the invention to bind or compete with a Therapeutic protein for binding to  
10 an anti-Therapeutic polypeptide antibody and/or anti-albumin antibody, various immunoassays known in the art can be used, including but not limited to, competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), “sandwich” immunoassays, immunoradiometric assays, gel diffusion precipitation reactions, immunodiffusion assays, in situ immunoassays  
15 (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is  
20 detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

In a preferred embodiment, where a binding partner (e.g., a receptor or a ligand) of a Therapeutic protein is identified, binding to that binding partner by an albumin fusion protein containing that Therapeutic protein as the Therapeutic protein portion of the fusion can be  
25 assayed, e.g., by means well-known in the art, such as, for example, reducing and non-reducing gel chromatography, protein affinity chromatography, and affinity blotting. See generally, Phizicky et al., Microbiol. Rev. 59:94-123 (1995). In another embodiment, the ability of physiological correlates of an albumin fusion protein of the present invention to  
30 bind to a substrate(s) of the Therapeutic polypeptide corresponding to the Therapeutic portion of the albumin fusion protein of the invention can be routinely assayed using techniques known in the art.

In an alternative embodiment, where the ability of an albumin fusion protein of the invention to multimerize is being evaluated, association with other components of the  
35 multimer can be assayed, e.g., by means well-known in the art, such as, for example, reducing and non-reducing gel chromatography, protein affinity chromatography, and affinity

blotting. See generally, Phizicky et al., *supra*.

In addition, assays described herein (see Examples and Table 1) and otherwise known in the art may routinely be applied to measure the ability of albumin fusion proteins of the present invention and fragments, variants and derivatives thereof to elicit biological activity and/or Therapeutic activity (either *in vitro* or *in vivo*) related to either the Therapeutic protein portion and/or albumin portion of the albumin fusion protein of the present invention. Other methods will be known to the skilled artisan and are within the scope of the invention.

### Albumin

As described above, an albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion or chemical conjugation.

The terms, human serum albumin (HSA) and human albumin (HA) are used interchangeably herein. The terms, "albumin" and "serum albumin" are broader, and encompass human serum albumin (and fragments and variants thereof) as well as albumin from other species (and fragments and variants thereof).

As used herein, "albumin" refers collectively to albumin protein or amino acid sequence, or an albumin fragment or variant, having one or more functional activities (e.g., biological activities) of albumin. In particular, "albumin" refers to human albumin or fragments thereof (see EP 201 239, EP 322 094 WO 97/24445, WO95/23857) especially the mature form of human albumin as shown in Figure 15 and SEQ ID NO:18, or albumin from other vertebrates or fragments thereof, or analogs or variants of these molecules or fragments thereof.

In preferred embodiments, the human serum albumin protein used in the albumin fusion proteins of the invention contains one or both of the following sets of point mutations with reference to SEQ ID NO:18: Leu-407 to Ala, Leu-408 to Val, Val-409 to Ala, and Arg-410 to Ala; or Arg-410 to A, Lys-413 to Gln, and Lys-414 to Gln (see, e.g., International Publication No. WO95/23857, hereby incorporated in its entirety by reference herein). In even more preferred embodiments, albumin fusion proteins of the invention that contain one or both of above-described sets of point mutations have improved stability/resistance to yeast Yap3p proteolytic cleavage, allowing increased production of recombinant albumin fusion proteins expressed in yeast host cells.

As used herein, a portion of albumin sufficient to prolong the therapeutic activity or shelf-life of the Therapeutic protein refers to a portion of albumin sufficient in length or structure to stabilize or prolong the therapeutic activity of the protein so that the shelf life of

the Therapeutic protein portion of the albumin fusion protein is prolonged or extended compared to the shelf-life in the non-fusion state. The albumin portion of the albumin fusion proteins may comprise the full length of the HA sequence as described above or as shown in Figure 15, or may include one or more fragments thereof that are capable of stabilizing or  
5 prolonging the therapeutic activity. Such fragments may be of 10 or more amino acids in length or may include about 15, 20, 25, 30, 50, or more contiguous amino acids from the HA sequence or may include part or all of specific domains of HA. For instance, one or more fragments of HA spanning the first two immunoglobulin-like domains may be used.

The albumin portion of the albumin fusion proteins of the invention may be a variant  
10 of normal HA. The Therapeutic protein portion of the albumin fusion proteins of the invention may also be variants of the Therapeutic proteins as described herein. The term "variants" includes insertions, deletions and substitutions, either conservative or non conservative, where such changes do not substantially alter one or more of the oncotic, useful ligand-binding and non-immunogenic properties of albumin, or the active site, or active  
15 domain which confers the therapeutic activities of the Therapeutic proteins.

In particular, the albumin fusion proteins of the invention may include naturally occurring polymorphic variants of human albumin and fragments of human albumin, for example those fragments disclosed in EP 322 094 (namely HA (Pn), where n is 369 to 419). The albumin may be derived from any vertebrate, especially any mammal, for example  
20 human, cow, sheep, or pig. Non-mammalian albumins include, but are not limited to, hen and salmon. The albumin portion of the albumin fusion protein may be from a different animal than the Therapeutic protein portion.

Generally speaking, an HA fragment or variant will be at least 100 amino acids long, preferably at least 150 amino acids long. The HA variant may consist of or alternatively  
25 comprise at least one whole domain of HA, for example domains 1 (amino acids 1-194 of SEQ ID NO:18), 2 (amino acids 195-387 of SEQ ID NO:18), 3 (amino acids 388-585 of SEQ ID NO:18), 1 + 2 (1-387 of SEQ ID NO:18), 2 + 3 (195-585 of SEQ ID NO:18) or 1 + 3 (amino acids 1-194 of SEQ ID NO:18 + amino acids 388-585 of SEQ ID NO:18). Each domain is itself made up of two homologous subdomains namely 1-105, 120-194, 195-291,  
30 316-387, 388-491 and 512-585, with flexible inter-subdomain linker regions comprising residues Lys106 to Glu119, Glu292 to Val315 and Glu492 to Ala511.

Preferably, the albumin portion of an albumin fusion protein of the invention comprises at least one subdomain or domain of HA or conservative modifications thereof. If the fusion is based on subdomains, some or all of the adjacent linker is preferably used to link  
35 to the Therapeutic protein moiety.

### Albumin Fusion Proteins

The present invention relates generally to albumin fusion proteins and methods of treating, preventing, or ameliorating diseases or disorders. As used herein, "albumin fusion protein" refers to a protein formed by the fusion of at least one molecule of albumin (or a fragment or variant thereof) to at least one molecule of a Therapeutic protein (or fragment or variant thereof). An albumin fusion protein of the invention comprises at least a fragment or variant of a Therapeutic protein and at least a fragment or variant of human serum albumin, which are associated with one another, preferably by genetic fusion (i.e., the albumin fusion protein is generated by translation of a nucleic acid in which a polynucleotide encoding all or a portion of a Therapeutic protein is joined in-frame with a polynucleotide encoding all or a portion of albumin) or chemical conjugation to one another. The Therapeutic protein and albumin protein, once part of the albumin fusion protein, may be referred to as a "portion", "region" or "moiety" of the albumin fusion protein.

In one embodiment, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein (e.g., as described in Table 1) and a serum albumin protein. In other embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment of a Therapeutic protein and a serum albumin protein. In other embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active variant of a Therapeutic protein and a serum albumin protein. In preferred embodiments, the serum albumin protein component of the albumin fusion protein is the mature portion of serum albumin.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein, and a biologically active and/or therapeutically active fragment of serum albumin. In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a Therapeutic protein and a biologically active and/or therapeutically active variant of serum albumin. In preferred embodiments, the Therapeutic protein portion of the albumin fusion protein is the mature portion of the Therapeutic protein.

In further embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, a biologically active and/or therapeutically active fragment or variant of a Therapeutic protein and a biologically active and/or therapeutically active fragment or variant of serum albumin. In preferred embodiments, the invention provides an albumin fusion protein comprising, or alternatively consisting of, the mature portion of a Therapeutic protein and the mature portion of serum albumin.

Preferably, the albumin fusion protein comprises HA as the N-terminal portion, and a

Therapeutic protein as the C-terminal portion. Alternatively, an albumin fusion protein comprising HA as the C-terminal portion, and a Therapeutic protein as the N-terminal portion may also be used.

In other embodiments, the albumin fusion protein has a Therapeutic protein fused to both the N-terminus and the C-terminus of albumin. In a preferred embodiment, the Therapeutic proteins fused at the N- and C- termini are the same Therapeutic proteins. In a preferred embodiment, the Therapeutic proteins fused at the N- and C- termini are different Therapeutic proteins. In another preferred embodiment, the Therapeutic proteins fused at the N- and C- termini are different Therapeutic proteins which may be used to treat or prevent the same disease, disorder, or condition (e.g. as listed in the "Preferred Indication Y" column of Table 1). In another preferred embodiment, the Therapeutic proteins fused at the N- and C-termini are different Therapeutic proteins which may be used to treat or prevent diseases or disorders (e.g. as listed in the "Preferred Indication Y" column of Table 1) which are known in the art to commonly occur in patients simultaneously.

In addition to albumin fusion protein in which the albumin portion is fused N-terminal and/or C-terminal of the Therapeutic protein portion, albumin fusion proteins of the invention may also be produced by inserting the Therapeutic protein or peptide of interest (e.g., Therapeutic protein X as disclosed in Table 1) into an internal region of HA. For instance, within the protein sequence of the HA molecule a number of loops or turns exist between the end and beginning of  $\alpha$ -helices, which are stabilized by disulphide bonds (see Figures 9-11). The loops, as determined from the crystal structure of HA (Fig. 13) (PDB identifiers 1AO6, 1BJ5, 1BKE, 1BM0, 1E7E to 1E7I and 1UOR) for the most part extend away from the body of the molecule. These loops are useful for the insertion, or internal fusion, of therapeutically active peptides, particularly those requiring a secondary structure to be functional, or Therapeutic proteins, to essentially generate an albumin molecule with specific biological activity.

Loops in human albumin structure into which peptides or polypeptides may be inserted to generate albumin fusion proteins of the invention include: Val54-Asn61, Thr76-Asp89, Ala92-Glu100, Gln170-Ala176, His247-Glu252, Glu266-Glu277, Glu280-His288, Ala362-Glu368, Lys439-Pro447, Val462-Lys475, Thr478-Pro486, and Lys560-Thr566. In more preferred embodiments, peptides or polypeptides are inserted into the Val54-Asn61, Gln170-Ala176, and/or Lys560-Thr566 loops of mature human albumin (SEQ ID NO:18).

Peptides to be inserted may be derived from either phage display or synthetic peptide libraries screened for specific biological activity or from the active portions of a molecule with the desired function. Additionally, random peptide libraries may be generated within particular loops or by insertions of randomized peptides into particular loops of the HA



molecule and in which all possible combinations of amino acids are represented.

Such library(s) could be generated on HA or domain fragments of HA by one of the following methods:

- (a) randomized mutation of amino acids within one or more peptide loops of HA or HA domain fragments. Either one, more or all the residues within a loop could be mutated in this manner (for example see Fig. 10a);
- (b) replacement of, or insertion into one or more loops of HA or HA domain fragments (*i.e.*, internal fusion) of a randomized peptide(s) of length  $X_n$  (where X is an amino acid and n is the number of residues (for example see Fig. 10b);
- (c) N-, C- or N- and C- terminal peptide/protein fusions in addition to (a) and/or (b).

The HA or HA domain fragment may also be made multifunctional by grafting the peptides derived from different screens of different loops against different targets into the same HA or HA domain fragment.

- 15 In preferred embodiments, peptides inserted into a loop of human serum albumin are peptide fragments or peptide variants of the Therapeutic proteins disclosed in Table 1. More particularly, the invention encompasses albumin fusion proteins which comprise peptide fragments or peptide variants at least 7 at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 35, or at least 40 amino acids in length inserted into a loop of human serum albumin. The invention also encompasses albumin fusion proteins which comprise peptide fragments or peptide variants at least 7 at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 35, or at least 40 amino acids fused to the N-terminus of human serum albumin. The invention also encompasses albumin fusion proteins which comprise peptide fragments or peptide variants at least 7 at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 35, or at least 40 amino acids fused to the C-terminus of human serum albumin.

- 25 Generally, the albumin fusion proteins of the invention may have one HA-derived region and one Therapeutic protein-derived region. Multiple regions of each protein, however, may be used to make an albumin fusion protein of the invention. Similarly, more than one Therapeutic protein may be used to make an albumin fusion protein of the invention.

- For instance, a Therapeutic protein may be fused to both the N- and C-terminal ends of the HA. In such a configuration, the Therapeutic protein portions may be the same or different Therapeutic protein molecules. The structure of bifunctional albumin fusion proteins may be represented as: X-HA-Y or Y-HA-X.

For example, an anti-BLyS™ scFv-HA-IFN $\alpha$ -2b fusion may be prepared to modulate

the immune response to IFN $\alpha$ -2b by anti-BLyS<sup>TM</sup> scFv. An alternative is making a bi (or even multi) functional dose of HA-fusions *e.g.* HA-IFN $\alpha$ -2b fusion mixed with HA-anti-BLyS<sup>TM</sup> scFv fusion or other HA-fusions in various ratio's depending on function, half-life etc.

5        Bi- or multi-functional albumin fusion proteins may also be prepared to target the Therapeutic protein portion of a fusion to a target organ or cell type via protein or peptide at the opposite terminus of HA.

10        As an alternative to the fusion of known therapeutic molecules, the peptides could be obtained by screening libraries constructed as fusions to the N-, C- or N- and C- termini of HA, or domain fragment of HA, of typically 6, 8, 12, 20 or 25 or X<sub>n</sub> (where X is an amino acid (aa) and n equals the number of residues) randomized amino acids, and in which all possible combinations of amino acids were represented. A particular advantage of this approach is that the peptides may be selected *in situ* on the HA molecule and the properties of the peptide would therefore be as selected for rather than, potentially, modified as might be  
15        the case for a peptide derived by any other method then being attached to HA.

20        Additionally, the albumin fusion proteins of the invention may include a linker peptide between the fused portions to provide greater physical separation between the moieties and thus maximize the accessibility of the Therapeutic protein portion, for instance, for binding to its cognate receptor. The linker peptide may consist of amino acids such that it is flexible or more rigid.

25        The linker sequence may be cleavable by a protease or chemically to yield the growth hormone related moiety. Preferably, the protease is one which is produced naturally by the host, for example the *S. cerevisiae* protease *kex2* or equivalent proteases.

30        Therefore, as described above, the albumin fusion proteins of the invention may have the following formula R1-L-R2; R2-L-R1; or R1-L-R2-L-R1, wherein R1 is at least one Therapeutic protein, peptide or polypeptide sequence, and not necessarily the same Therapeutic protein, L is a linker and R2 is a serum albumin sequence.

35        In preferred embodiments, Albumin fusion proteins of the invention comprising a Therapeutic protein have extended shelf life compared to the shelf life the same Therapeutic protein when not fused to albumin. Shelf-life typically refers to the time period over which the therapeutic activity of a Therapeutic protein in solution or in some other storage formulation, is stable without undue loss of therapeutic activity. Many of the Therapeutic proteins are highly labile in their unfused state. As described below, the typical shelf-life of these Therapeutic proteins is markedly prolonged upon incorporation into the albumin fusion protein of the invention.

Albumin fusion proteins of the invention with "prolonged" or "extended" shelf-life exhibit greater therapeutic activity relative to a standard that has been subjected to the same storage and handling conditions. The standard may be the unfused full-length Therapeutic protein. When the Therapeutic protein portion of the albumin fusion protein is an analog, a  
5 variant, or is otherwise altered or does not include the complete sequence for that protein, the prolongation of therapeutic activity may alternatively be compared to the unfused equivalent of that analog, variant, altered peptide or incomplete sequence. As an example, an albumin fusion protein of the invention may retain greater than about 100% of the therapeutic activity, or greater than about 105%, 110%, 120%, 130%, 150% or 200% of the therapeutic activity  
10 of a standard when subjected to the same storage and handling conditions as the standard when compared at a given time point.

Shelf-life may also be assessed in terms of therapeutic activity remaining after storage, normalized to therapeutic activity when storage began. Albumin fusion proteins of the invention with prolonged or extended shelf-life as exhibited by prolonged or extended  
15 therapeutic activity may retain greater than about 50% of the therapeutic activity, about 60%, 70%, 80%, or 90% or more of the therapeutic activity of the equivalent unfused Therapeutic protein when subjected to the same conditions. For example, as discussed in Example 1, an albumin fusion protein of the invention comprising hGH fused to the full length HA sequence may retain about 80% or more of its original activity in solution for periods of up to 5 weeks  
20 or more under various temperature conditions.

### *Expression of Fusion Proteins*

The albumin fusion proteins of the invention may be produced as recombinant molecules by secretion from yeast, a microorganism such as a bacterium, or a human or  
25 animal cell line. Preferably, the polypeptide is secreted from the host cells. We have found that, by fusing the hGH coding sequence to the HA coding sequence, either to the 5' end or 3' end, it is possible to secrete the albumin fusion protein from yeast without the requirement for a yeast-derived pro sequence. This was surprising, as other workers have found that a yeast derived pro sequence was needed for efficient secretion of hGH in yeast.

For example, Hiramatsu *et al.* (Appl Environ Microbiol 56:2125 (1990); Appl Environ Microbiol 57:2052 (1991)) found that the N-terminal portion of the pro sequence in the *Mucor pusillus* rennin pre-pro leader was important. Other authors, using the MF $\alpha$ -1 signal, have  
30 always included the MF $\alpha$ -1 pro sequence when secreting hGH. The pro sequences were believed to assist in the folding of the hGH by acting as an intramolecular chaperone. The  
35 present invention shows that HA or fragments of HA can perform a similar function.

Hence, a particular embodiment of the invention comprises a DNA construct encoding a signal sequence effective for directing secretion in yeast, particularly a yeast-derived signal sequence (especially one which is homologous to the yeast host), and the fused molecule of the first aspect of the invention, there being no yeast-derived pro sequence  
5 between the signal and the mature polypeptide.

The *Saccharomyces cerevisiae* invertase signal is a preferred example of a yeast-derived signal sequence.

Conjugates of the kind prepared by Poznansky *et al.*, (FEBS Lett. 239:18 (1988)), in which separately-prepared polypeptides are joined by chemical cross-linking, are not  
10 contemplated.

The present invention also includes a cell, preferably a yeast cell transformed to express an albumin fusion protein of the invention. In addition to the transformed host cells themselves, the present invention also contemplates a culture of those cells, preferably a monoclonal (clonally homogeneous) culture, or a culture derived from a monoclonal culture,  
15 in a nutrient medium. If the polypeptide is secreted, the medium will contain the polypeptide, with the cells, or without the cells if they have been filtered or centrifuged away. Many expression systems are known and may be used, including bacteria (for example *E. coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*, *Kluyveromyces lactis* and *Pichia pastoris*, filamentous fungi (for example *Aspergillus*), plant cells, animal cells and  
20 insect cells.

Preferred yeast strains to be used in the production of albumin fusion proteins are D88, DXY1 and BXP10. D88 [*leu2-3*, *leu2-122*, *can1*, *pra1*, *ubc4*] is a derivative of parent strain AH22<sup>his+</sup> (also known as DB1; see, e.g., Sleep *et al.* Biotechnology 8:42-46 (1990)). The strain contains a *leu2* mutation which allows for auxotrophic selection of 2 micron-based  
25 plasmids that contain the LEU2 gene. D88 also exhibits a derepression of PRB1 in glucose excess. The PRB1 promoter is normally controlled by two checkpoints that monitor glucose levels and growth stage. The promoter is activated in wild type yeast upon glucose depletion and entry into stationary phase. Strain D88 exhibits the repression by glucose but maintains the induction upon entry into stationary phase. The PRA1 gene encodes a yeast vacuolar protease, YscA endoprotease A, that is localized in the ER. The UBC4 gene is in the ubiquitination pathway and is involved in targeting short lived and abnormal proteins for ubiquitin dependant degradation. Isolation of this *ubc4* mutation was found to increase the copy number of an expression plasmid in the cell and cause an increased level of expression of a desired protein expressed from the plasmid (see, e.g., International Publication No.  
30 WO99/00504, hereby incorporated in its entirety by reference herein).

DXY1, a derivative of D88, has the following genotype: [*leu2-3*, *leu2-122*, *can1*,

*pral*, *ubc4*, *ura3::yap3*]. In addition to the mutations isolated in D88, this strain also has a knockout of the YAP3 protease. This protease causes cleavage of mostly di-basic residues (RR, RK, KR, KK) but can also promote cleavage at single basic residues in proteins. Isolation of this *yap3* mutation resulted in higher levels of full length HSA production (see, e.g., U.S. Patent No. 5,965,386, hereby incorporated in its entirety by reference herein, and Kerry-Williams et al., Yeast 14:161-169 (1998), hereby incorporated in their entireties by reference herein).

BXP10 has the following genotype: *leu2-3*, *leu2-122*, *can1*, *pral*, *ubc4*, *ura3*, *yap3::URA3*, *lys2*, *hsp150::LYS2*, *pmt1::URA3*. In addition to the mutations isolated in DXY1, this strain also has a knockout of the PMT1 gene and the HSP150 gene. The PMT1 gene is a member of the evolutionarily conserved family of dolichyl-phosphate-D-mannose protein O-mannosyltransferases (Pmts). The transmembrane topology of Pmt1p suggests that it is an integral membrane protein of the endoplasmic reticulum with a role in O-linked glycosylation. This mutation serves to reduce/eliminate O-linked glycosylation of HSA fusions (see, e.g., International Publication No. WO00/44772, hereby incorporated in its entirety by reference herein). Studies revealed that the Hsp150 protein is inefficiently separated from rHA by ion exchange chromatography. The mutation in the HSP150 gene removes a potential contaminant that has proven difficult to remove by standard purification techniques. See, e.g., U.S. Patent No. 5,783,423, hereby incorporated in its entirety by reference herein.

The desired protein is produced in conventional ways, for example from a coding sequence inserted in the host chromosome or on a free plasmid. The yeasts are transformed with a coding sequence for the desired protein in any of the usual ways, for example electroporation. Methods for transformation of yeast by electroporation are disclosed in Becker & Guarente (1990) *Methods Enzymol.* 194, 182.

Successfully transformed cells, i.e., cells that contain a DNA construct of the present invention, can be identified by well known techniques. For example, cells resulting from the introduction of an expression construct can be grown to produce the desired polypeptide. Cells can be harvested and lysed and their DNA content examined for the presence of the DNA using a method such as that described by Southern (1975) *J. Mol. Biol.* 98, 503 or Berent et al. (1985) *Biotech.* 3, 208. Alternatively, the presence of the protein in the supernatant can be detected using antibodies.

Useful yeast plasmid vectors include pRS403-406 and pRS413-416 and are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Plasmids pRS403, pRS404, pRS405 and pRS406 are Yeast Integrating plasmids (YIps) and incorporate the yeast selectable markers HIS3, 7RP1, LEU2 and URA3. Plasmids

pRS413-416 are Yeast Centromere plasmids (Ycps).

Preferred vectors for making albumin fusion proteins for expression in yeast include pPPC0005, pScCHSA, pScNHSA, and pC4:HSA which are described in detail in Example 2. Figure 4 shows a map of the pPPC0005 plasmid that can be used as the base vector into which polynucleotides encoding Therapeutic proteins may be cloned to form HA-fusions. It contains a *PRB1 S. cerevisiae* promoter (PRB1p), a Fusion leader sequence (FL), DNA encoding HA (rHA) and an *ADH1 S. cerevisiae* terminator sequence. The sequence of the fusion leader sequence consists of the first 19 amino acids of the signal peptide of human serum albumin (SEQ ID NO:29) and the last five amino acids of the mating factor alpha 1 promoter (SLDKR, see EP-A-387 319 which is hereby incorporated by reference in its entirety.

The plasmids, pPPC0005, pScCHSA, pScNHSA, and pC4:HSA were deposited on April 11, 2001 at the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209, and given accession numbers ATCC \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_, respectively. Another vector useful for expressing an albumin fusion protein in yeast the pSAC35 vector which is described in Sleep *et al.*, BioTechnology 8:42 (1990) which is hereby incorporated by reference in its entirety.

A variety of methods have been developed to operably link DNA to vectors via complementary cohesive termini. For instance, complementary homopolymer tracts can be added to the DNA segment to be inserted to the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, generated by endonuclease restriction digestion, is treated with bacteriophage T4 DNA polymerase or E. coli DNA polymerase I, enzymes that remove protruding,  $\gamma$ -single-stranded termini with their 3' 5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerizing activities.

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a large molar excess of linker molecules in the presence of an enzyme that is able to catalyze the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including International Biotechnologies Inc,

New Haven, CT, USA.

A desirable way to modify the DNA in accordance with the invention, if, for example, HA variants are to be prepared, is to use the polymerase chain reaction as disclosed by Saiki *et al.* (1988) *Science* 239, 487-491. In this method the DNA to be enzymatically amplified is flanked by two specific oligonucleotide primers which themselves become incorporated into the amplified DNA. The specific primers may contain restriction endonuclease recognition sites which can be used for cloning into expression vectors using methods known in the art.

Exemplary genera of yeast contemplated to be useful in the practice of the present invention as hosts for expressing the albumin fusion proteins are *Pichia* (formerly classified as *Hansenula*), *Saccharomyces*, *Kluyveromyces*, *Aspergillus*, *Candida*, *Torulopsis*, *Torulaspora*, *Schizosaccharomyces*, *Citeromyces*, *Pachysolen*, *Zygosaccharomyces*, *Debaromyces*, *Trichoderma*, *Cephalosporium*, *Humicola*, *Mucor*, *Neurospora*, *Yarrowia*, *Metschnikowia*, *Rhodospiridium*, *Leucosporidium*, *Botryosphaeria*, *Sporidiobolus*, *Endomycopsis*, and the like. Preferred genera are those selected from the group consisting of *Saccharomyces*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia* and *Torulaspora*. Examples of *Saccharomyces* spp. are *S. cerevisiae*, *S. italicus* and *S. rouxii*.

Examples of *Kluyveromyces* spp. are *K. fragilis*, *K. lactis* and *K. marxianus*. A suitable *Torulaspora* species is *T. delbrueckii*. Examples of *Pichia* (*Hansenula*) spp. are *P. angusta* (formerly *H. polymorpha*), *P. anomala* (formerly *H. anomala*) and *P. pastoris*. Methods for the transformation of *S. cerevisiae* are taught generally in EP 251 744, EP 258 067 and WO 90/01063, all of which are incorporated herein by reference.

Preferred exemplary species of *Saccharomyces* include *S. cerevisiae*, *S. italicus*, *S. diastaticus*, and *Zygosaccharomyces rouxii*. Preferred exemplary species of *Kluyveromyces* include *K. fragilis* and *K. lactis*. Preferred exemplary species of *Hansenula* include *H. polymorpha* (now *Pichia angusta*), *H. anomala* (now *Pichia anomala*), and *Pichia capsulata*. Additional preferred exemplary species of *Pichia* include *P. pastoris*. Preferred exemplary species of *Aspergillus* include *A. niger* and *A. nidulans*. Preferred exemplary species of *Yarrowia* include *Y. lipolytica*. Many preferred yeast species are available from the ATCC. For example, the following preferred yeast species are available from the ATCC and are useful in the expression of albumin fusion proteins: *Saccharomyces cerevisiae* Hansen, teleomorph strain BY4743 *yap3* mutant (ATCC Accession No. 4022731); *Saccharomyces cerevisiae* Hansen, teleomorph strain BY4743 *hsp150* mutant (ATCC Accession No. 4021266); *Saccharomyces cerevisiae* Hansen, teleomorph strain BY4743 *pmt1* mutant (ATCC Accession No. 4023792); *Saccharomyces cerevisiae* Hansen, teleomorph (ATCC Accession Nos. 20626; 44773; 44774; and 62995); *Saccharomyces diastaticus* Andrews et Gilliland ex van der Walt, teleomorph (ATCC Accession No. 62987); *Kluyveromyces lactis*

(Dombrowski) van der Walt, teleomorph (ATCC Accession No. 76492); *Pichia angusta* (Teunisson et al.) Kurtzman, teleomorph deposited as *Hansenula polymorpha* de Morais et Maia, teleomorph (ATCC Accession No. 26012); *Aspergillus niger* van Tieghem, anamorph (ATCC Accession No. 9029); *Aspergillus niger* van Tieghem, anamorph (ATCC Accession No. 16404); *Aspergillus nidulans* (Eidam) Winter, anamorph (ATCC Accession No. 48756); and *Yarrowia lipolytica* (Wickerham et al.) van der Walt et von Arx, teleomorph (ATCC Accession No. 201847).

Suitable promoters for *S. cerevisiae* include those associated with the PGKI gene, GAL1 or GAL10 genes, CYC1, PHO5, TRP1, ADHI, ADH2, the genes for glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, triose phosphate isomerase, phosphoglucose isomerase, glucokinase, alpha-mating factor pheromone, [a mating factor pheromone], the PRB1 promoter, the GUT2 promoter, the GPDI promoter, and hybrid promoters involving hybrids of parts of 5' regulatory regions with parts of 5' regulatory regions of other promoters or with upstream activation sites (e.g. the promoter of EP-A-258 067).

Convenient regulatable promoters for use in *Schizosaccharomyces pombe* are the thiamine-repressible promoter from the nmt gene as described by Maundrell (1990) *J. Biol. Chem.* 265, 10857-10864 and the glucose repressible jbp1 gene promoter as described by Hoffman & Winston (1990) *Genetics* 124, 807-816.

Methods of transforming *Pichia* for expression of foreign genes are taught in, for example, Cregg *et al.* (1993), and various Phillips patents (e.g. US 4 857 467, incorporated herein by reference), and *Pichia* expression kits are commercially available from Invitrogen BV, Leek, Netherlands, and Invitrogen Corp., San Diego, California. Suitable promoters include AOX1 and AOX2. Gleeson *et al.* (1986) *J. Gen. Microbiol.* 132, 3459-3465 include information on *Hansenula* vectors and transformation, suitable promoters being MOX1 and FMD1; whilst EP 361 991, Fleer *et al.* (1991) and other- publications from Rhone-Poulenc Rorer teach how to express foreign proteins in *Kluyveromyces* spp., a suitable promoter being PGKI.

The transcription termination signal is preferably the 3' flanking sequence of a eukaryotic gene which contains proper signals for transcription termination and polyadenylation. Suitable 3' flanking sequences may, for example, be those of the gene naturally linked to the expression control sequence used, i.e. may correspond to the promoter. Alternatively, they may be different in which case the termination signal of the *S. cerevisiae* ADHI gene is preferred.

The desired albumin fusion protein may be initially expressed with a secretion leader sequence, which may be any leader effective in the yeast chosen. Leaders useful in *S.*



cerevisiae include that from the mating factor  $\alpha$  polypeptide (MF  $\alpha$ -1) and the hybrid leaders of EP-A-387 319. Such leaders (or signals) are cleaved by the yeast before the mature albumin is released into the surrounding medium. Further such leaders include those of *S. cerevisiae* invertase (SUC2) disclosed in JP 62-096086 (granted as 911036516), acid phosphatase (PH05), the pre-sequence of MF $\alpha$ -1,  $\alpha$  glucanase (BGL2) and killer toxin; *S. diastaticus* glucoamylase II; *S. carlsbergensis*  $\alpha$ -galactosidase (MEL1); *K. lactis* killer toxin; and *Candida glucoamylase*.

#### Additional Methods of Recombinant and Synthetic Production of Albumin Fusion Proteins

The present invention also relates to vectors containing a polynucleotide encoding an albumin fusion protein of the present invention, host cells, and the production of albumin fusion proteins by synthetic and recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides encoding albumin fusion proteins of the invention may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp*, *phoA* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418, glutamine synthase, or neomycin resistance for eukaryotic cell culture, and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces*

and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as CHO, COS, NSO, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Preferred expression vectors for use in yeast systems include, but are not limited to pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, pPIC9K, and PAO815 (all available from Invitrogen, Carlsbad, CA). Other suitable vectors will be readily apparent to the skilled artisan.

In one embodiment, polynucleotides encoding an albumin fusion protein of the invention may be fused to signal sequences which will direct the localization of a protein of the invention to particular compartments of a prokaryotic or eukaryotic cell and/or direct the secretion of a protein of the invention from a prokaryotic or eukaryotic cell. For example, in *E. coli*, one may wish to direct the expression of the protein to the periplasmic space. Examples of signal sequences or proteins (or fragments thereof) to which the albumin fusion proteins of the invention may be fused in order to direct the expression of the polypeptide to the periplasmic space of bacteria include, but are not limited to, the *pelB* signal sequence, the maltose binding protein (MBP) signal sequence, MBP, the *ompA* signal sequence, the signal sequence of the periplasmic *E. coli* heat-labile enterotoxin B-subunit, and the signal sequence of alkaline phosphatase. Several vectors are commercially available for the construction of fusion proteins which will direct the localization of a protein, such as the pMAL series of vectors (particularly the pMAL-p series) available from New England Biolabs. In a specific embodiment, polynucleotides albumin fusion proteins of the invention may be fused to the *pelB* pectate lyase signal sequence to increase the efficiency of expression and purification of such polypeptides in Gram-negative bacteria. See, U.S. Patent Nos. 5,576,195 and 5,846,818, the contents of which are herein incorporated by reference in their entireties.

Examples of signal peptides that may be fused to an albumin fusion protein of the invention in order to direct its secretion in mammalian cells include, but are not limited to, the MPIF-1 signal sequence (e.g., amino acids 1-21 of GenBank Accession number AAB51134), the stanniocalcin signal sequence (MLQNSAVLLLLVISASA, SEQ ID NO:34),

and a consensus signal sequence (MPTWAWWLFLVLLLALWAPARG, SEQ ID NO:35). A suitable signal sequence that may be used in conjunction with baculoviral expression systems is the gp67 signal sequence (e.g., amino acids 1-19 of GenBank Accession Number AAA72759).

5        Vectors which use glutamine synthase (GS) or DHFR as the selectable markers can be amplified in the presence of the drugs methionine sulfoximine or methotrexate, respectively. An advantage of glutamine synthase based vectors are the availability of cell lines (e.g., the murine myeloma cell line, NSO) which are glutamine synthase negative. Glutamine synthase expression systems can also function in glutamine synthase expressing cells (e.g., Chinese  
10    Hamster Ovary (CHO) cells) by providing additional inhibitor to prevent the functioning of the endogenous gene. A glutamine synthase expression system and components thereof are detailed in PCT publications: WO87/04462; WO86/05807; WO89/01036; WO89/10404; and WO91/06657, which are hereby incorporated in their entireties by reference herein. Additionally, glutamine synthase expression vectors can be obtained from Lonza Biologics,  
15    Inc. (Portsmouth, NH). Expression and production of monoclonal antibodies using a GS expression system in murine myeloma cells is described in Bebbington *et al.*, *Bio/technology* 10:169(1992) and in Biblia and Robinson *Biotechnol. Prog.* 11:1 (1995) which are herein incorporated by reference.

      The present invention also relates to host cells containing the above-described vector  
20    constructs described herein, and additionally encompasses host cells containing nucleotide sequences of the invention that are operably associated with one or more heterologous control regions (e.g., promoter and/or enhancer) using techniques known of in the art. The host cell can be a higher eukaryotic cell, such as a mammalian cell (e.g., a human derived cell), or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a  
25    bacterial cell. A host strain may be chosen which modulates the expression of the inserted gene sequences, or modifies and processes the gene product in the specific fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus expression of the genetically engineered polypeptide may be controlled. Furthermore, different host cells have characteristics and specific mechanisms for the translational and post-  
30    translational processing and modification (e.g., phosphorylation, cleavage) of proteins. Appropriate cell lines can be chosen to ensure the desired modifications and processing of the foreign protein expressed.

      Introduction of the nucleic acids and nucleic acid constructs of the invention into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated  
35    transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as

Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

5 In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., the coding sequence corresponding to a Therapeutic protein may be replaced with an albumin fusion protein corresponding to the Therapeutic protein), and/or to include genetic material (e.g., heterologous polynucleotide sequences such as for example, an albumin fusion protein of the invention corresponding to the Therapeutic protein may be included). The genetic material operably associated with the endogenous polynucleotide may activate, alter, and/or amplify endogenous polynucleotides.

10 In addition, techniques known in the art may be used to operably associate heterologous polynucleotides (e.g., polynucleotides encoding an albumin protein, or a fragment or variant thereof) and/or heterologous control regions (e.g., promoter and/or enhancer) with endogenous polynucleotide sequences encoding a Therapeutic protein via homologous recombination (see, e.g., US Patent Number 5,641,670, issued June 24, 1997; International Publication Number WO 96/29411; International Publication Number WO 94/12650; Koller *et al.*, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 (1989); and Zijlstra *et al.*, *Nature* 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

15 Albumin fusion proteins of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, hydrophobic charge interaction chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

20 In preferred embodiments the albumin fusion proteins of the invention are purified using Anion Exchange Chromatography including, but not limited to, chromatography on Q-sepharose, DEAE sepharose, poros HQ, poros DEAE, Toyopearl Q, Toyopearl QAE, Toyopearl DEAE, Resource/Source Q and DEAE, Fractogel Q and DEAE columns.

25 In specific embodiments the albumin fusion proteins of the invention are purified using Cation Exchange Chromatography including, but not limited to, SP-sepharose, CM sepharose, poros HS, poros CM, Toyopearl SP, Toyopearl CM, Resource/Source S and CM, Fractogel S and CM columns and their equivalents and comparables.

In specific embodiments the albumin fusion proteins of the invention are purified using Hydrophobic Interaction Chromatography including, but not limited to, Phenyl, Butyl, Methyl, Octyl, Hexyl-sepharose, poros Phenyl, Butyl, Methyl, Octyl, Hexyl, Toyopearl Phenyl, Butyl, Methyl, Octyl, Hexyl Resource/Source Phenyl, Butyl, Methyl, Octyl, Hexyl, Fractogel Phenyl, Butyl, Methyl, Octyl, Hexyl columns and their equivalents and comparables.

In specific embodiments the albumin fusion proteins of the invention are purified using Size Exclusion Chromatography including, but not limited to, sepharose S100, S200, S300, superdex resin columns and their equivalents and comparables.

In specific embodiments the albumin fusion proteins of the invention are purified using Affinity Chromatography including, but not limited to, Mimetic Dye affinity, peptide affinity and antibody affinity columns that are selective for either the HSA or the "fusion target" molecules.

In preferred embodiments albumin fusion proteins of the invention are purified using one or more Chromatography methods listed above. In other preferred embodiments, albumin fusion proteins of the invention are purified using one or more of the following Chromatography columns, Q sepharose FF column, SP Sepharose FF column, Q Sepharose High Performance Column, Blue Sepharose FF column, Blue Column, Phenyl Sepharose FF column, DEAE Sepharose FF, or Methyl Column.

Additionally, albumin fusion proteins of the invention may be purified using the process described in International Publication No. WO00/44772 which is herein incorporated by reference in its entirety. One of skill in the art could easily modify the process described therein for use in the purification of albumin fusion proteins of the invention.

Albumin fusion proteins of the present invention may be recovered from: products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, albumin fusion proteins of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In one embodiment, the yeast *Pichia pastoris* is used to express albumin fusion proteins of the invention in a eukaryotic system. *Pichia pastoris* is a methylotrophic yeast which can metabolize methanol as its sole carbon source. A main step in the methanol metabolism pathway is the oxidation of methanol to formaldehyde using O<sub>2</sub>. This reaction is catalyzed by the enzyme alcohol oxidase. In order to metabolize methanol as its sole carbon source, *Pichia pastoris* must generate high levels of alcohol oxidase due, in part, to the relatively low affinity of alcohol oxidase for O<sub>2</sub>. Consequently, in a growth medium depending on methanol as a main carbon source, the promoter region of one of the two alcohol oxidase genes (*AOX1*) is highly active. In the presence of methanol, alcohol oxidase produced from the *AOX1* gene comprises up to approximately 30% of the total soluble protein in *Pichia pastoris*. See Ellis, S.B., *et al.*, *Mol. Cell. Biol.* 5:1111-21 (1985); Koutz, P.J., *et al.*, *Yeast* 5:167-77 (1989); Tschopp, J.F., *et al.*, *Nucl. Acids Res.* 15:3859-76 (1987). Thus, a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, under the transcriptional regulation of all or part of the *AOX1* regulatory sequence is expressed at exceptionally high levels in *Pichia* yeast grown in the presence of methanol.

In one example, the plasmid vector pPIC9K is used to express DNA encoding an albumin fusion protein of the invention, as set forth herein, in a *Pichea* yeast system essentially as described in "*Pichia* Protocols: Methods in Molecular Biology," D.R. Higgins and J. Cregg, eds. The Humana Press, Totowa, NJ, 1998. This expression vector allows expression and secretion of a polypeptide of the invention by virtue of the strong *AOX1* promoter linked to the *Pichia pastoris* alkaline phosphatase (PHO) secretory signal peptide (i.e., leader) located upstream of a multiple cloning site.

Many other yeast vectors could be used in place of pPIC9K, such as, pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalpha, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, and PAO815, as one skilled in the art would readily appreciate, as long as the proposed expression construct provides appropriately located signals for transcription, translation, secretion (if desired), and the like, including an in-frame AUG as required.

In another embodiment, high-level expression of a heterologous coding sequence, such as, for example, a polynucleotide encoding an albumin fusion protein of the present invention, may be achieved by cloning the heterologous polynucleotide of the invention into an expression vector such as, for example, pGAPZ or pGAPZalpha, and growing the yeast culture in the absence of methanol.

In addition, albumin fusion proteins of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, *Proteins: Structures and*

Molecular Principles, W.H. Freeman & Co., N.Y., and Hunkapiller et al., *Nature*, 310:105-111 (1984)). For example, a polypeptide corresponding to a fragment of a polypeptide can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the polypeptide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid,  $\gamma$ -Abu,  $\epsilon$ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine,  $\beta$ -alanine, fluoro-amino acids, designer amino acids such as  $\beta$ -methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

The invention encompasses albumin fusion proteins of the present invention which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease,  $\text{NaBH}_4$ ; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The albumin fusion proteins may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin; and examples of suitable radioactive material include iodine ( $^{121}\text{I}$ ,  $^{123}\text{I}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{111}\text{In}$ ,  $^{112}\text{In}$ ,  $^{113\text{m}}\text{In}$ ,  $^{115\text{m}}\text{In}$ ), technetium ( $^{99}\text{Tc}$ ,  $^{99\text{m}}\text{Tc}$ ),

thallium ( $^{201}\text{Ti}$ ), gallium ( $^{68}\text{Ga}$ ,  $^{67}\text{Ga}$ ), palladium ( $^{103}\text{Pd}$ ), molybdenum ( $^{99}\text{Mo}$ ), xenon ( $^{133}\text{Xe}$ ), fluorine ( $^{18}\text{F}$ ),  $^{153}\text{Sm}$ ,  $^{177}\text{Lu}$ ,  $^{159}\text{Gd}$ ,  $^{149}\text{Pm}$ ,  $^{140}\text{La}$ ,  $^{175}\text{Yb}$ ,  $^{166}\text{Ho}$ ,  $^{90}\text{Y}$ ,  $^{47}\text{Sc}$ ,  $^{186}\text{Re}$ ,  $^{188}\text{Re}$ ,  $^{142}\text{Pr}$ ,  $^{105}\text{Rh}$ , and  $^{97}\text{Ru}$ .

In specific embodiments, albumin fusion proteins of the present invention or  
5 fragments or variants thereof are attached to macrocyclic chelators that associate with  
radiometal ions, including but not limited to,  $^{177}\text{Lu}$ ,  $^{90}\text{Y}$ ,  $^{166}\text{Ho}$ , and  $^{153}\text{Sm}$ , to polypeptides.  
In a preferred embodiment, the radiometal ion associated with the macrocyclic chelators is  
 $^{111}\text{In}$ . In another preferred embodiment, the radiometal ion associated with the macrocyclic  
chelator is  $^{90}\text{Y}$ . In specific embodiments, the macrocyclic chelator is 1,4,7,10-  
10 tetraazacyclododecane-N,N',N'',N'''-tetraacetic acid (DOTA). In other specific  
embodiments, DOTA is attached to an antibody of the invention or fragment thereof via linker  
molecule. Examples of linker molecules useful for conjugating DOTA to a polypeptide are  
commonly known in the art - see, for example, DeNardo et al., Clin Cancer Res. 4(10):2483-  
90 (1998); Peterson et al., Bioconjug. Chem. 10(4):553-7 (1999); and Zimmerman et al.,  
15 Nucl. Med. Biol. 26(8):943-50 (1999); which are hereby incorporated by reference in their  
entirety.

As mentioned, the albumin fusion proteins of the invention may be modified by either  
natural processes, such as post-translational processing, or by chemical modification  
techniques which are well known in the art. It will be appreciated that the same type of  
20 modification may be present in the same or varying degrees at several sites in a given  
polypeptide. Polypeptides of the invention may be branched, for example, as a result of  
ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and  
branched cyclic polypeptides may result from posttranslation natural processes or may be  
made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation,  
25 amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent  
attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid  
derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide  
bond formation, demethylation, formation of covalent cross-links, formation of cysteine,  
formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor  
30 formation, hydroxylation, iodination, methylation, myristylation, oxidation, pegylation,  
proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation,  
transfer-RNA mediated addition of amino acids to proteins such as arginylation, and  
ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR  
PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York  
35 (1993); POST-TRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C.  
Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth. Enzymol.



182:626-646 (1990); Rattan et al., *Ann. N.Y. Acad. Sci.* 663:48-62 (1992)).

Albumin fusion proteins of the invention and antibodies that bind a Therapeutic protein or fragments or variants thereof can be fused to marker sequences, such as a peptide to facilitate purification. In preferred embodiments, the marker amino acid sequence is a  
5 hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., *Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Other peptide tags useful for purification include, but are not limited to, the "HA" tag, which corresponds to an  
10 epitope derived from the influenza hemagglutinin protein (Wilson et al., *Cell* 37:767 (1984)) and the "flag" tag.

Further, an albumin fusion protein of the invention may be conjugated to a therapeutic moiety such as a cytotoxin, e.g., a cytostatic or cytotoxic agent, a therapeutic agent or a radioactive metal ion, e.g., alpha-emitters such as, for example, <sup>213</sup>Bi. A cytotoxin or  
15 cytotoxic agent includes any agent that is detrimental to cells. Examples include paclitaxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, teniposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs  
20 thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis- dichlorodiamine platinum (II) (DDP) cisplatin),  
25 anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the therapeutic agent or drug moiety is not to be construed as limited to classical  
30 chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, an apoptotic agent, e.g., TNF-alpha, TNF-beta, AIM I  
35 (See, International Publication No. WO 97/33899), AIM II (See, International Publication No. WO 97/34911), Fas Ligand (Takahashi *et al.*, *Int. Immunol.*, 6:1567-1574 (1994)),

VEGI (See, International Publication No. WO 99/23105), a thrombotic agent or an anti-angiogenic agent, e.g., angiostatin or endostatin; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Techniques for conjugating such therapeutic moiety to proteins (e.g., albumin fusion proteins) are well known in the art.

Albumin fusion proteins may also be attached to solid supports, which are particularly useful for immunoassays or purification of polypeptides that are bound by, that bind to, or associate with albumin fusion proteins of the invention. Such solid supports include, but are not limited to, glass, cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene.

Albumin fusion proteins, with or without a therapeutic moiety conjugated to it, administered alone or in combination with cytotoxic factor(s) and/or cytokine(s) can be used as a therapeutic.

Also provided by the invention are chemically modified derivatives of the albumin fusion proteins of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The albumin fusion proteins may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a Therapeutic protein or analog). For example, the polyethylene glycol may have an average molecular weight of about 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10,000, 10,500, 11,000, 11,500, 12,000, 12,500, 13,000, 13,500, 14,000, 14,500, 15,000, 15,500, 16,000, 16,500, 17,000, 17,500, 18,000, 18,500, 19,000, 19,500, 20,000, 25,000, 30,000, 35,000, 40,000, 45,000, 50,000, 55,000, 60,000, 65,000, 70,000, 75,000,

80,000, 85,000, 90,000, 95,000, or 100,000 kDa.

As noted above, the polyethylene glycol may have a branched structure. Branched polyethylene glycols are described, for example, in U.S. Patent No. 5,643,575; Morpurgo *et al.*, *Appl. Biochem. Biotechnol.* 56:59-72 (1996); Vorobjev *et al.*, *Nucleosides Nucleotides* 18:2745-2750 (1999); and Caliceti *et al.*, *Bioconjug. Chem.* 10:638-646 (1999), the disclosures of each of which are incorporated herein by reference.

The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those skilled in the art, such as, for example, the method disclosed in EP 0 401 384 (coupling PEG to G-CSF), herein incorporated by reference; see also Malik *et al.*, *Exp. Hematol.* 20:1028-1035 (1992), reporting pegylation of GM-CSF using tresyl chloride. For example, polyethylene glycol may be covalently bound through amino acid residues via reactive group, such as a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

As suggested above, polyethylene glycol may be attached to proteins via linkage to any of a number of amino acid residues. For example, polyethylene glycol can be linked to proteins via covalent bonds to lysine, histidine, aspartic acid, glutamic acid, or cysteine residues. One or more reaction chemistries may be employed to attach polyethylene glycol to specific amino acid residues (e.g., lysine, histidine, aspartic acid, glutamic acid, or cysteine) of the protein or to more than one type of amino acid residue (e.g., lysine, histidine, aspartic acid, glutamic acid, cysteine and combinations thereof) of the protein.

One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus modification may

be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

As indicated above, pegylation of the albumin fusion proteins of the invention may be accomplished by any number of means. For example, polyethylene glycol may be attached to the albumin fusion protein either directly or by an intervening linker. Linkerless systems for attaching polyethylene glycol to proteins are described in Delgado et al., Crit. Rev. Thera. Drug Carrier Sys. 9:249-304 (1992); Francis et al., Intern. J. of Hematol. 68:1-18 (1998); U.S. Patent No. 4,002,531; U.S. Patent No. 5,349,052; WO 95/06058; and WO 98/32466, the disclosures of each of which are incorporated herein by reference.

One system for attaching polyethylene glycol directly to amino acid residues of proteins without an intervening linker employs tresylated MPEG, which is produced by the modification of monmethoxy polyethylene glycol (MPEG) using tresylchloride ( $\text{ClSO}_2\text{CH}_2\text{CF}_3$ ). Upon reaction of protein with tresylated MPEG, polyethylene glycol is directly attached to amine groups of the protein. Thus, the invention includes protein-polyethylene glycol conjugates produced by reacting proteins of the invention with a polyethylene glycol molecule having a 2,2,2-trifluoroethane sulphonyl group.

Polyethylene glycol can also be attached to proteins using a number of different intervening linkers. For example, U.S. Patent No. 5,612,460, the entire disclosure of which is incorporated herein by reference, discloses urethane linkers for connecting polyethylene glycol to proteins. Protein-polyethylene glycol conjugates wherein the polyethylene glycol is attached to the protein by a linker can also be produced by reaction of proteins with compounds such as MPEG-succinimidylsuccinate, MPEG activated with 1,1'-carbonyldiimidazole, MPEG-2,4,5-trichloropenylcarbonate, MPEG-p-nitrophenolcarbonate, and various MPEG-succinate derivatives. A number of additional polyethylene glycol derivatives and reaction chemistries for attaching polyethylene glycol to proteins are described in International Publication No. WO 98/32466, the entire disclosure of which is incorporated herein by reference. Pegylated protein products produced using the reaction chemistries set out herein are included within the scope of the invention.

The number of polyethylene glycol moieties attached to each albumin fusion protein of the invention (i.e., the degree of substitution) may also vary. For example, the pegylated proteins of the invention may be linked, on average, to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20, or more polyethylene glycol molecules. Similarly, the average degree of substitution within ranges such as 1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9, 8-10, 9-11, 10-12, 11-13, 12-14,

13-15, 14-16, 15-17, 16-18, 17-19, or 18-20 polyethylene glycol moieties per protein molecule. Methods for determining the degree of substitution are discussed, for example, in Delgado et al., Crit. Rev. Thera. Drug Carrier Sys. 9:249-304 (1992).

5 The polypeptides of the invention can be recovered and purified from chemical synthesis and recombinant cell cultures by standard methods which include, but are not limited to, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is  
10 employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and/or purification.

The presence and quantity of albumin fusion proteins of the invention may be determined using ELISA, a well known immunoassay known in the art. In one ELISA  
15 protocol that would be useful for detecting/quantifying albumin fusion proteins of the invention, comprises the steps of coating an ELISA plate with an anti-human serum albumin antibody, blocking the plate to prevent non-specific binding, washing the ELISA plate, adding a solution containing the albumin fusion protein of the invention (at one or more different concentrations), adding a secondary anti-Therapeutic protein specific antibody  
20 coupled to a detectable label (as described herein or otherwise known in the art), and detecting the presence of the secondary antibody. In an alternate version of this protocol, the ELISA plate might be coated with the anti-Therapeutic protein specific antibody and the labeled secondary reagent might be the anti-human albumin specific antibody.

## 25 Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

30 The polynucleotides of the present invention are useful to produce the albumin fusion proteins of the invention. As described in more detail below, polynucleotides of the invention (encoding albumin fusion proteins) may be used in recombinant DNA methods useful in genetic engineering to make cells, cell lines, or tissues that express the albumin fusion protein encoded by the polynucleotides encoding albumin fusion proteins of the invention.

35 Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a

means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell. Additional non-limiting examples of gene therapy methods encompassed by the present invention are more thoroughly described elsewhere herein (see, e.g., the sections labeled

5 "Gene Therapy", and Examples 17 and 18).

### Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

10 Albumin fusion proteins of the invention are useful to provide immunological probes for differential identification of the tissue(s) (e.g., immunohistochemistry assays such as, for example, ABC immunoperoxidase (Hsu et al., J. Histochem. Cytochem. 29:577-580 (1981)) or cell type(s) (e.g., immunocytochemistry assays).

Albumin fusion proteins can be used to assay levels of polypeptides in a biological

15 sample using classical immunohistological methods known to those of skill in the art (e.g., see Jalkanen, et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, et al., J. Cell. Biol. 105:3087-3096 (1987)). Other methods useful for detecting protein gene expression include immunoassays; such as the enzyme linked immunosorbent assay (ELISA) and the

20 radioimmunoassay (RIA). Suitable assay labels are known in the art and include enzyme labels, such as, glucose oxidase; radioisotopes, such as iodine ( $^{131}\text{I}$ ,  $^{125}\text{I}$ ,  $^{123}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{115\text{m}}\text{In}$ ,  $^{113\text{m}}\text{In}$ ,  $^{112}\text{In}$ ,  $^{111}\text{In}$ ), and technetium ( $^{99\text{m}}\text{Tc}$ ,  $^{99\text{m}}\text{Tc}$ ), thallium ( $^{201}\text{Tl}$ ), gallium ( $^{68}\text{Ga}$ ,  $^{67}\text{Ga}$ ), palladium ( $^{103}\text{Pd}$ ), molybdenum ( $^{99}\text{Mo}$ ), xenon ( $^{133}\text{Xe}$ ), fluorine ( $^{18}\text{F}$ ),  $^{153}\text{Sm}$ ,  $^{177}\text{Lu}$ ,  $^{159}\text{Gd}$ ,  $^{149}\text{Pm}$ ,  $^{140}\text{La}$ ,  $^{175}\text{Yb}$ ,  $^{166}\text{Ho}$ ,  $^{90}\text{Y}$ ,  $^{47}\text{Sc}$ ,  $^{186}\text{Re}$ ,  $^{188}\text{Re}$ ,  $^{142}\text{Pr}$ ,  $^{105}\text{Rh}$ ,  $^{97}\text{Ru}$ ; luminescent labels, such as luminol; and fluorescent labels, such as

25 fluorescein and rhodamine, and biotin.

Albumin fusion proteins of the invention can also be detected *in vivo* by imaging. Labels or markers for *in vivo* imaging of protein include those detectable by X-radiography, nuclear magnetic resonance (NMR) or electron spin relaxation (ESR). For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation

30 but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the albumin fusion protein by labeling of nutrients given to a cell line expressing the albumin fusion protein of the invention.

An albumin fusion protein which has been labeled with an appropriate detectable

35 imaging moiety, such as a radioisotope (for example,  $^{131}\text{I}$ ,  $^{112}\text{In}$ ,  $^{99\text{m}}\text{Tc}$ , ( $^{131}\text{I}$ ,  $^{125}\text{I}$ ,  $^{123}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{115\text{m}}\text{In}$ ,  $^{113\text{m}}\text{In}$ ,  $^{112}\text{In}$ ,  $^{111}\text{In}$ ), and technetium

(<sup>99</sup>Tc, <sup>99m</sup>Tc), thallium (<sup>201</sup>Tl), gallium (<sup>68</sup>Ga, <sup>67</sup>Ga), palladium (<sup>103</sup>Pd), molybdenum (<sup>99</sup>Mo), xenon (<sup>133</sup>Xe), fluorine (<sup>18</sup>F, <sup>153</sup>Sm, <sup>177</sup>Lu, <sup>159</sup>Gd, <sup>149</sup>Pm, <sup>140</sup>La, <sup>175</sup>Yb, <sup>166</sup>Ho, <sup>90</sup>Y, <sup>47</sup>Sc, <sup>186</sup>Re, <sup>188</sup>Re, <sup>142</sup>Pr, <sup>105</sup>Rh, <sup>97</sup>Ru), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously or intraperitoneally) into the mammal to be examined for immune system disorder. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of <sup>99m</sup>Tc. The labeled albumin fusion protein will then preferentially accumulate at locations in the body (e.g., organs, cells, extracellular spaces or matrices) where one or more receptors, ligands or substrates (corresponding to that of the Therapeutic protein used to make the albumin fusion protein of the invention) are located. Alternatively, in the case where the albumin fusion protein comprises at least a fragment or variant of a Therapeutic antibody, the labeled albumin fusion protein will then preferentially accumulate at the locations in the body (e.g., organs, cells, extracellular spaces or matrices) where the polypeptides/epitopes corresponding to those bound by the Therapeutic antibody (used to make the albumin fusion protein of the invention) are located. *In vivo* tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments" (Chapter 13 in *Tumor Imaging: The Radiochemical Detection of Cancer*, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982)). The protocols described therein could easily be modified by one of skill in the art for use with the albumin fusion proteins of the invention.

In one embodiment, the invention provides a method for the specific delivery of albumin fusion proteins of the invention to cells by administering albumin fusion proteins of the invention (e.g., polypeptides encoded by polynucleotides encoding albumin fusion proteins of the invention and/or antibodies) that are associated with heterologous polypeptides or nucleic acids. In one example, the invention provides a method for delivering a Therapeutic protein into the targeted cell. In another example, the invention provides a method for delivering a single stranded nucleic acid (e.g., antisense or ribozymes) or double stranded nucleic acid (e.g., DNA that can integrate into the cell's genome or replicate episomally and that can be transcribed) into the targeted cell.

In another embodiment, the invention provides a method for the specific destruction of cells (e.g., the destruction of tumor cells) by administering albumin fusion proteins of the invention in association with toxins or cytotoxic prodrugs.

By "toxin" is meant one or more compounds that bind and activate endogenous cytotoxic effector systems, radioisotopes, holotoxins, modified toxins, catalytic subunits of

toxins, or any molecules or enzymes not normally present in or on the surface of a cell that under defined conditions cause the cell's death. Toxins that may be used according to the methods of the invention include, but are not limited to, radioisotopes known in the art, compounds such as, for example, antibodies (or complement fixing containing portions thereof) that bind an inherent or induced endogenous cytotoxic effector system, thymidine kinase, endonuclease, RNase, alpha toxin, ricin, abrin, *Pseudomonas* exotoxin A, diphtheria toxin, saporin, momordin, gelonin, pokeweed antiviral protein, alpha-sarcin and cholera toxin. "Toxin" also includes a cytostatic or cytocidal agent, a therapeutic agent or a radioactive metal ion, e.g., alpha-emitters such as, for example,  $^{213}\text{Bi}$ , or other radioisotopes such as, for example,  $^{103}\text{Pd}$ ,  $^{133}\text{Xe}$ ,  $^{131}\text{I}$ ,  $^{68}\text{Ge}$ ,  $^{57}\text{Co}$ ,  $^{65}\text{Zn}$ ,  $^{85}\text{Sr}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{90}\text{Y}$ ,  $^{153}\text{Sm}$ ,  $^{153}\text{Gd}$ ,  $^{169}\text{Yb}$ ,  $^{51}\text{Cr}$ ,  $^{54}\text{Mn}$ ,  $^{75}\text{Se}$ ,  $^{113}\text{Sn}$ ,  $^{90}\text{Yttrium}$ ,  $^{117}\text{Tin}$ ,  $^{186}\text{Rhenium}$ ,  $^{166}\text{Holmium}$ , and  $^{188}\text{Rhenium}$ ; luminescent labels, such as luminol; and fluorescent labels, such as fluorescein and rhodamine, and biotin. In a specific embodiment, the invention provides a method for the specific destruction of cells (e.g., the destruction of tumor cells) by administering polypeptides of the invention or antibodies of the invention in association with the radioisotope  $^{90}\text{Y}$ . In another specific embodiment, the invention provides a method for the specific destruction of cells (e.g., the destruction of tumor cells) by administering polypeptides of the invention or antibodies of the invention in association with the radioisotope  $^{111}\text{In}$ . In a further specific embodiment, the invention provides a method for the specific destruction of cells (e.g., the destruction of tumor cells) by administering polypeptides of the invention or antibodies of the invention in association with the radioisotope  $^{131}\text{I}$ .

Techniques known in the art may be applied to label polypeptides of the invention. Such techniques include, but are not limited to, the use of bifunctional conjugating agents (see e.g., U.S. Patent Nos. 5,756,065; 5,714,631; 5,696,239; 5,652,361; 5,505,931; 5,489,425; 5,435,990; 5,428,139; 5,342,604; 5,274,119; 4,994,560; and 5,808,003; the contents of each of which are hereby incorporated by reference in its entirety).

The albumin fusion proteins of the present invention are useful for diagnosis, treatment, prevention and/or prognosis of various disorders in mammals, preferably humans. Such disorders include, but are not limited to, those described herein under the section heading "Biological Activities," below.

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression level of a certain polypeptide in cells or body fluid of an individual using an albumin fusion protein of the invention; and (b) comparing the assayed polypeptide expression level with a standard polypeptide expression level, whereby an increase or decrease in the assayed polypeptide expression level compared to the standard expression



level is indicative of a disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Moreover, albumin fusion proteins of the present invention can be used to treat or prevent diseases or conditions such as, for example, neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders, renal disorders, proliferative disorders, and/or cancerous diseases and conditions. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B, SOD, catalase, DNA repair proteins), to inhibit the activity of a polypeptide (e.g., an oncogene or tumor suppressor), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth inhibition, enhancement of the immune response to proliferative cells or tissues).

In particular, albumin fusion proteins comprising of at least a fragment or variant of a Therapeutic antibody can also be used to treat disease (as described *supra*, and elsewhere herein). For example, administration of an albumin fusion protein comprising of at least a fragment or variant of a Therapeutic antibody can bind, and/or neutralize the polypeptide to which the Therapeutic antibody used to make the albumin fusion protein immunospecifically binds, and/or reduce overproduction of the polypeptide to which the Therapeutic antibody used to make the albumin fusion protein immunospecifically binds. Similarly, administration of an albumin fusion protein comprising of at least a fragment or variant of a Therapeutic antibody can activate the polypeptide to which the Therapeutic antibody used to make the albumin fusion protein immunospecifically binds, by binding to the polypeptide bound to a membrane (receptor).

At the very least, the albumin fusion proteins of the invention of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Albumin fusion proteins of the invention can also be used to raise antibodies, which in turn may be used to measure protein expression of the Therapeutic protein, albumin protein, and/or the albumin fusion protein of the invention from a recombinant cell, as a way of assessing transformation

of the host cell, or in a biological sample. Moreover, the albumin fusion proteins of the present invention can be used to test the biological activities described herein.

### *Diagnostic Assays*

5       The compounds of the present invention are useful for diagnosis, treatment, prevention and/or prognosis of various disorders in mammals, preferably humans. Such disorders include, but are not limited to, those described for each Therapeutic protein in the corresponding row of Table 1 and herein under the section headings "Immune Activity," "Blood Related Disorders," "Hyperproliferative Disorders," "Renal Disorders,"  
10 "Cardiovascular Disorders," "Respiratory Disorders," "Anti-Angiogenesis Activity," "Diseases at the Cellular Level," "Wound Healing and Epithelial Cell Proliferation," "Neural Activity and Neurological Diseases," "Endocrine Disorders," "Reproductive System Disorders," "Infectious Disease," "Regeneration," and/or "Gastrointestinal Disorders," *infra*.

For a number of disorders, substantially altered (increased or decreased) levels of  
15 gene expression can be detected in tissues, cells or bodily fluids (e.g., sera, plasma, urine, semen, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" gene expression level, that is, the expression level in tissues or bodily fluids from an individual not having the disorder. Thus, the invention provides a diagnostic method useful during diagnosis of a disorder, which involves measuring the expression level of the  
20 gene encoding a polypeptide in tissues, cells or body fluid from an individual and comparing the measured gene expression level with a standard gene expression level, whereby an increase or decrease in the gene expression level(s) compared to the standard is indicative of a disorder. These diagnostic assays may be performed *in vivo* or *in vitro*, such as, for example, on blood samples, biopsy tissue or autopsy tissue.

25       The present invention is also useful as a prognostic indicator, whereby patients exhibiting enhanced or depressed gene expression will experience a worse clinical outcome

By "assaying the expression level of the gene encoding a polypeptide" is intended qualitatively or quantitatively measuring or estimating the level of a particular polypeptide (e.g. a polypeptide corresponding to a Therapeutic protein disclosed in Table 1) or the level of  
30 the mRNA encoding the polypeptide of the invention in a first biological sample either directly (e.g., by determining or estimating absolute protein level or mRNA level) or relatively (e.g., by comparing to the polypeptide level or mRNA level in a second biological sample). Preferably, the polypeptide expression level or mRNA level in the first biological sample is measured or estimated and compared to a standard polypeptide level or mRNA level, the  
35 standard being taken from a second biological sample obtained from an individual not having the disorder or being determined by averaging levels from a population of individuals not

having the disorder. As will be appreciated in the art, once a standard polypeptide level or mRNA level is known, it can be used repeatedly as a standard for comparison.

By "biological sample" is intended any biological sample obtained from an individual, cell line, tissue culture, or other source containing polypeptides of the invention (including portions thereof) or mRNA. As indicated, biological samples include body fluids (such as sera, plasma, urine, synovial fluid and spinal fluid) and tissue sources found to express the full length or fragments thereof of a polypeptide or mRNA. Methods for obtaining tissue biopsies and body fluids from mammals are well known in the art. Where the biological sample is to include mRNA, a tissue biopsy is the preferred source.

Total cellular RNA can be isolated from a biological sample using any suitable technique such as the single-step guanidinium-thiocyanate-phenol-chloroform method described in Chomczynski and Sacchi, *Anal. Biochem.* 162:156-159 (1987). Levels of mRNA encoding the polypeptides of the invention are then assayed using any appropriate method. These include Northern blot analysis, S1 nuclease mapping, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR), and reverse transcription in combination with the ligase chain reaction (RT-LCR).

The present invention also relates to diagnostic assays such as quantitative and diagnostic assays for detecting levels of polypeptides that bind to, are bound by, or associate with albumin fusion proteins of the invention, in a biological sample (e.g., cells and tissues), including determination of normal and abnormal levels of polypeptides. Thus, for instance, a diagnostic assay in accordance with the invention for detecting abnormal expression of polypeptides that bind to, are bound by, or associate with albumin fusion proteins compared to normal control tissue samples may be used to detect the presence of tumors. Assay techniques that can be used to determine levels of a polypeptide that bind to, are bound by, or associate with albumin fusion proteins of the present invention in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Assaying polypeptide levels in a biological sample can occur using any art-known method.

Assaying polypeptide levels in a biological sample can occur using a variety of techniques. For example, polypeptide expression in tissues can be studied with classical immunohistological methods (Jalkanen et al., *J. Cell. Biol.* 101:976-985 (1985); Jalkanen, M., et al., *J. Cell. Biol.* 105:3087-3096 (1987)). Other methods useful for detecting polypeptide gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine ( $^{125}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium

(<sup>112</sup>In), and technetium (<sup>99m</sup>Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

The tissue or cell type to be analyzed will generally include those which are known, or suspected, to express the gene of interest (such as, for example, cancer). The protein isolation methods employed herein may, for example, be such as those described in Harlow and Lane (Harlow, E. and Lane, D., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), which is incorporated herein by reference in its entirety. The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells that could be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of the gene.

For example, albumin fusion proteins may be used to quantitatively or qualitatively detect the presence of polypeptides that bind to, are bound by, or associate with albumin fusion proteins of the present invention. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled albumin fusion protein coupled with light microscopic, flow cytometric, or fluorimetric detection.

In a preferred embodiment, albumin fusion proteins comprising at least a fragment or variant of an antibody that immunospecifically binds at least a Therapeutic protein disclosed herein (e.g., the Therapeutic proteins disclosed in Table 1) or otherwise known in the art may be used to quantitatively or qualitatively detect the presence of gene products or conserved variants or peptide fragments thereof. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody coupled with light microscopic, flow cytometric, or fluorimetric detection.

The albumin fusion proteins of the present invention may, additionally, be employed histologically, as in immunofluorescence, immunoelectron microscopy or non-immunological assays, for in situ detection of polypeptides that bind to, are bound by, or associate with an albumin fusion protein of the present invention. In situ detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody or polypeptide of the present invention. The albumin fusion proteins are preferably applied by overlaying the labeled albumin fusion proteins onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the polypeptides that bind to, are bound by, or associate with albumin fusion proteins, but also its distribution in the examined tissue. Using the present invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such in situ detection.

Immunoassays and non-immunoassays that detect polypeptides that bind to, are

bound by, or associate with albumin fusion proteins will typically comprise incubating a sample, such as a biological fluid, a tissue extract, freshly harvested cells, or lysates of cells which have been incubated in cell culture, in the presence of a detectably labeled antibody capable of binding gene products or conserved variants or peptide fragments thereof, and  
5 detecting the bound antibody by any of a number of techniques well-known in the art.

The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support which is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with  
10 suitable buffers followed by treatment with the detectably labeled albumin fusion protein of the invention. The solid phase support may then be washed with the buffer a second time to remove unbound antibody or polypeptide. Optionally the antibody is subsequently labeled. The amount of bound label on solid support may then be detected by conventional means.

By "solid phase support or carrier" is intended any support capable of binding a polypeptide (e.g., an albumin fusion protein, or polypeptide that binds, is bound by, or  
15 associates with an albumin fusion protein of the invention.) Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the  
20 coupled molecule is capable of binding to a polypeptide. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use  
25 of routine experimentation.

The binding activity of a given lot of albumin fusion protein may be determined according to well known methods. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation.

In addition to assaying polypeptide levels in a biological sample obtained from an  
30 individual, polypeptide can also be detected *in vivo* by imaging. For example, in one embodiment of the invention, albumin fusion proteins of the invention are used to image diseased or neoplastic cells.

Labels or markers for *in vivo* imaging of albumin fusion proteins of the invention include those detectable by X-radiography, NMR, MRI, CAT-scans or ESR. For  
35 X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and

ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the albumin fusion protein by labeling of nutrients of a cell line (or bacterial or yeast strain) engineered.

5 Additionally, albumin fusion proteins of the invention whose presence can be detected, can be administered. For example, albumin fusion proteins of the invention labeled with a radio-opaque or other appropriate compound can be administered and visualized *in vivo*, as discussed, above for labeled antibodies. Further, such polypeptides can be utilized for *in vitro* diagnostic procedures.

10 A polypeptide-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example,  $^{131}\text{I}$ ,  $^{112}\text{In}$ ,  $^{99\text{m}}\text{Tc}$ ), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously or intraperitoneally) into the mammal to be examined for a disorder. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce  
15 diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of  $^{99\text{m}}\text{Tc}$ . The labeled albumin fusion protein will then preferentially accumulate at the locations in the body which contain a polypeptide or other substance that binds to, is bound by or associates with an albumin fusion protein of the present invention. *In vivo* tumor imaging is described in S.W.  
20 Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments" (Chapter 13 in *Tumor Imaging: The Radiochemical Detection of Cancer*, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982)).

One of the ways in which an albumin fusion protein of the present invention can be detectably labeled is by linking the same to a reporter enzyme and using the linked product in  
25 an enzyme immunoassay (EIA) (Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2:1-7, Microbiological Associates Quarterly Publication, Walkersville, MD); Voller et al., *J. Clin. Pathol.* 31:507-520 (1978); Butler, J.E., *Meth. Enzymol.* 73:482-523 (1981); Maggio, E. (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, FL.; Ishikawa, E. et al., (eds.), 1981, Enzyme Immunoassay,  
30 Kigaku Shoin, Tokyo). The reporter enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Reporter enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid  
35 isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose

oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. Additionally, the detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the reporter enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Albumin fusion proteins may also be radiolabelled and used in any of a variety of other immunoassays. For example, by radioactively labeling the albumin fusion proteins, it is possible to use the albumin fusion proteins in a radioimmunoassay (RIA) (see, for example, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986, which is incorporated by reference herein). The radioactive isotope can be detected by means including, but not limited to, a gamma counter, a scintillation counter, or autoradiography.

It is also possible to label the albumin fusion proteins with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, ophthaldehyde and fluorescamine.

The albumin fusion protein can also be detectably labeled using fluorescence emitting metals such as  $^{152}\text{Eu}$ , or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The albumin fusion proteins can also be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged albumin fusion protein is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, therromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label albumin fusion proteins of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in, which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

### Transgenic Organisms

Transgenic organisms that express the albumin fusion proteins of the invention are also included in the invention. Transgenic organisms are genetically modified organisms into which recombinant, exogenous or cloned genetic material has been transferred. Such genetic material is often referred to as a transgene. The nucleic acid sequence of the transgene may include one or more transcriptional regulatory sequences and other nucleic acid sequences such as introns, that may be necessary for optimal expression and secretion of the encoded protein. The transgene may be designed to direct the expression of the encoded protein in a manner that facilitates its recovery from the organism or from a product produced by the organism, *e.g.* from the milk, blood, urine, eggs, hair or seeds of the organism. The transgene may consist of nucleic acid sequences derived from the genome of the same species or of a different species than the species of the target animal. The transgene may be integrated either at a locus of a genome where that particular nucleic acid sequence is not otherwise normally found or at the normal locus for the transgene.

The term "germ cell line transgenic organism" refers to a transgenic organism in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability of the transgenic organism to transfer the genetic information to offspring. If such offspring in fact possess some or all of that alteration or genetic information, then they too are transgenic organisms. The alteration or genetic information may be foreign to the species of organism to which the recipient belongs, foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.

A transgenic organism may be a transgenic animal or a transgenic plant. Transgenic animals can be produced by a variety of different methods including transfection, electroporation, microinjection, gene targeting in embryonic stem cells and recombinant viral and retroviral infection (*see, e.g.*, U.S. Patent No. 4,736,866; U.S. Patent No. 5,602,307; Mullins *et al.* (1993) *Hypertension* 22(4):630-633; Brenin *et al.* (1997) *Surg. Oncol.* 6(2):99-110; Tuan (ed.), *Recombinant Gene Expression Protocols*, Methods in Molecular Biology No. 62, Humana Press (1997)). The method of introduction of nucleic acid fragments into recombination competent mammalian cells can be by any method which favors co-transformation of multiple nucleic acid molecules. Detailed procedures for producing transgenic animals are readily available to one skilled in the art, including the disclosures in U.S. Patent No. 5,489,743 and U.S. Patent No. 5,602,307.

A number of recombinant or transgenic mice have been produced, including those which express an activated oncogene sequence (U.S. Patent No. 4,736,866); express simian SV40 T-antigen (U.S. Patent No. 5,728,915); lack the expression of interferon regulatory



factor 1 (IRF-1) (U.S. Patent No. 5,731,490); exhibit dopaminergic dysfunction (U.S. Patent No. 5,723,719); express at least one human gene which participates in blood pressure control (U.S. Patent No. 5,731,489); display greater similarity to the conditions existing in naturally occurring Alzheimer's disease (U.S. Patent No. 5,720,936); have a reduced capacity to mediate cellular adhesion (U.S. Patent No. 5,602,307); possess a bovine growth hormone gene (Clutter *et al.* (1996) Genetics 143(4):1753-1760); or, are capable of generating a fully human antibody response (McCarthy (1997) The Lancet 349(9049):405).

While mice and rats remain the animals of choice for most transgenic experimentation, in some instances it is preferable or even necessary to use alternative animal species. Transgenic procedures have been successfully utilized in a variety of non-murine animals, including sheep, goats, pigs, dogs, cats, monkeys, chimpanzees, hamsters, rabbits, cows and guinea pigs (*see, e.g.,* Kim *et al.* (1997) Mol. Reprod. Dev. 46(4):515-526; Houdebine (1995) Reprod. Nutr. Dev. 35(6):609-617; Petters (1994) Reprod. Fertil. Dev. 6(5):643-645; Schnieke *et al.* (1997) Science 278(5346):2130-2133; and Amoah (1997) J. Animal Science 75(2):578-585).

To direct the secretion of the transgene-encoded protein of the invention into the milk of transgenic mammals, it may be put under the control of a promoter that is preferentially activated in mammary epithelial cells. Promoters that control the genes encoding milk proteins are preferred, for example the promoter for casein, beta lactoglobulin, whey acid protein, or lactalbumin (*see, e.g.,* DiTullio (1992) BioTechnology 10:74-77; Clark *et al.* (1989) BioTechnology 7:487-492; Gorton *et al.* (1987) BioTechnology 5:1183-1187; and Soulier *et al.* (1992) FEBS Letts. 297:13). The transgenic mammals of choice would produce large volumes of milk and have long lactating periods, for example goats, cows, camels or sheep.

An albumin fusion protein of the invention can also be expressed in a transgenic plant, *e.g.* a plant in which the DNA transgene is inserted into the nuclear or plastidic genome. Plant transformation procedures used to introduce foreign nucleic acids into plant cells or protoplasts are known in the art (*e.g., see* Example 19). See, in general, Methods in Enzymology Vol. 153 ("Recombinant DNA Part D") 1987, Wu and Grossman Eds., Academic Press and European Patent Application EP 693554. Methods for generation of genetically engineered plants are further described in US Patent No. 5,283,184, US Patent No. 5,482,852, and European Patent Application EP 693 554, all of which are hereby incorporated by reference.

### **Pharmaceutical or Therapeutic Compositions**

The albumin fusion proteins of the invention or formulations thereof may be

administered by any conventional method including parenteral (*e.g.* subcutaneous or intramuscular) injection or intravenous infusion. The treatment may consist of a single dose or a plurality of doses over a period of time.

While it is possible for an albumin fusion protein of the invention to be administered alone, it is preferable to present it as a pharmaceutical formulation, together with one or more acceptable carriers. The carrier(s) must be "acceptable" in the sense of being compatible with the albumin fusion protein and not deleterious to the recipients thereof. Typically, the carriers will be water or saline which will be sterile and pyrogen free. Albumin fusion proteins of the invention are particularly well suited to formulation in aqueous carriers such as sterile pyrogen free water, saline or other isotonic solutions because of their extended shelf-life in solution. For instance, pharmaceutical compositions of the invention may be formulated well in advance in aqueous form, for instance, weeks or months or longer time periods before being dispensed.

For example, wherein the Therapeutic protein is hGH, EPO, alpha-IFN or beta-IFN, formulations containing the albumin fusion protein may be prepared taking into account the extended shelf-life of the albumin fusion protein in aqueous formulations. As exhibited in Table 2, most Therapeutic proteins are unstable with short shelf-lives after formulation with an aqueous carrier. As discussed above, the shelf-life of many of these Therapeutic proteins are markedly increased or prolonged after fusion to HA.

**Table 2**

Protein	Tradename, Manufacturer	Route	Formulation	Storage Conditions of Non-Fusion Protein
Interferon, alpha-2a	Roferon-A, Hoffmann-LaRoche	sc im	sol_n (vial or pre-filled syringe)	4-8°C
Interferon, alpha-2b	Intron-A, Schering Plough	iv sc im	sol_n; powder + dil.	4-8°C (all preps, before and after dilution)
COMBO Interferon alpha-2b + Ribavirin	Rebetron (Intron-A + Rebetol) Schering Plough	po + sc	Rebetol capsule + Intron-A injection	
Interferon, Alphacon-1	Infergen Amgen	sc	sol_n	4-8°C
Interferon, alpha-n1, Lymphoblastoid	Wellferon, Wellcome	sc im	sol_n (with albumin _as stablizer_)	4-8°C

Protein	Tradename, Manufacturer	Route	Formulation	Storage Conditions of Non-Fusion Protein
Interferon, beta-1a	Avonex, Biogen	im	powder + dil. (with albumin)	4-8°C (before and after dilution) (Use within 3-6h of reconstitution)
	Rebif, Ares-Serono (Europe only)	sc	sol_n, in pre-filled syringe	
Interferon, beta-1b	Betaseron, Chiron (Europe: Betaferon)	sc	powder + dil. (with albumin)	4-8°C (before and after dilution) (Use within 3h of reconstitution) Single use vials.
Interferon, Gamma-1b	Actimmune, InterMune Pharmaceuticals	sc		4-8°C (before and after dilution) (Use within 3h of reconstitution).
Growth Hormone (somatropin)	Genotropin, Pharmacia Upjohn		powder/dil cartridges (single or multi-use); single use MiniQuick injector	4-8°C (before and after dilution); single use MiniQuick Delivery Device should be refrigerated until use.
	Humatrope, Eli Lilly	sc im	powder + dil. (Vial or pen cartridge)	4-8°C (before and after dilution) (Use vials within 25h, cartridges within 28d, of reconstitution).
	Norditropin, Novo Nordisk Pharmaceuticals			
	Nutropin, Genentech	sc	powder + dil.	4-8°C (stable for 14d after dil_n) (all preps, before and after dilution)
	Nutropin AQ, Genentech	sc	sol_n	4-8°C (Stable for 28 d after 1st use)
	Nutropin Depot, Genentech	sc	microsphere suspension as powder + dil.	4-8°C Single use pkges. Dose 1-2x/month (ProLease micro-encapsulation technol.)
	Saizen, (Serono)	sc im	powder + dil.	Powder should be stored at Rm Temp_. After reconstitution store 4- 8°C for up to 14d.

Protein	Tradename, Manufacturer	Route	Formulation	Storage Conditions of Non-Fusion Protein
	Serostim, Serono			Powder should be stored at Rm Temp. After reconstitution store in 4- 8°C for up to 14d.
hGH, with N-term. Met (somatrem)	Protropin, Genentech	sc im	powder + dil.	4-8°C (all preps, before and after dilution)
Erythropoietin (Epoetin alfa)	Epogen, Amgen	iv sc	sol_n	4-8°C (use within 21d of first use) (Single & multi-dose vials)
	Procrit, Amgen	iv sc	sol_n	4-8°C (use within 21d of first use) (Single & multi-dose vials)

In instances where aerosol administration is appropriate, the albumin fusion proteins of the invention can be formulated as aerosols using standard procedures. The term "aerosol" includes any gas-borne suspended phase of an albumin fusion protein of the instant invention which is capable of being inhaled into the bronchioles or nasal passages. Specifically, aerosol includes a gas-borne suspension of droplets of an albumin fusion protein of the instant invention, as may be produced in a metered dose inhaler or nebulizer, or in a mist sprayer. Aerosol also includes a dry powder composition of a compound of the instant invention suspended in air or other carrier gas, which may be delivered by insufflation from an inhaler device, for example. See Ganderton & Jones, *Drug Delivery to the Respiratory Tract*, Ellis Horwood (19 87); Gonda (1990) *Critical Reviews in Therapeutic Drug Carrier Systems* 6:273-313; and Raeburn *et al.*, (1992) *Pharmacol. Toxicol. Methods* 27:143-159.

The formulations of the invention are also typically non-immunogenic, in part, because of the use of the components of the albumin fusion protein being derived from the proper species. For instance, for human use, both the Therapeutic protein and albumin portions of the albumin fusion protein will typically be human. In some cases, wherein either component is non human-derived, that component may be humanized by substitution of key amino acids so that specific epitopes appear to the human immune system to be human in nature rather than foreign.

The formulations may conveniently be presented in unit dosage form and may be prepared by any of the methods well known in the art of pharmacy. Such methods include the step of bringing into association the albumin fusion protein with the carrier that constitutes one or more accessory ingredients. In general the formulations are prepared by uniformly and  
5 intimately bringing into association the active ingredient with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product.

Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation appropriate for the intended recipient; and aqueous and  
10 non-aqueous sterile suspensions which may include suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example sealed ampules, vials or syringes, and may be stored in a freeze-dried (lyophilised) condition requiring only the addition of the sterile liquid carrier, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be  
15 prepared from sterile powders. Dosage formulations may contain the Therapeutic protein portion at a lower molar concentration or lower dosage compared to the non-fused standard formulation for the Therapeutic protein given the extended serum half-life exhibited by many of the albumin fusion proteins of the invention.

As an example, when an albumin fusion protein of the invention comprises growth  
20 hormone as one or more of the Therapeutic protein regions, the dosage form can be calculated on the basis of the potency of the albumin fusion protein relative to the potency of hGH, while taking into account the prolonged serum half-life and shelf-life of the albumin fusion proteins compared to that of native hGH. Growth hormone is typically administered at 0.3 to 30.0 IU/kg/week, for example 0.9 to 12.0 IU/kg/week, given in three or seven divided doses  
25 for a year or more. In an albumin fusion protein consisting of full length HA fused to full length GH, an equivalent dose in terms of units would represent a greater weight of agent but the dosage frequency can be reduced, for example to twice a week, once a week or less.

Formulations or compositions of the invention may be packaged together with, or included in a kit with, instructions or a package insert referring to the extended shelf-life of  
30 the albumin fusion protein component. For instance, such instructions or package inserts may address recommended storage conditions, such as time, temperature and light, taking into account the extended or prolonged shelf-life of the albumin fusion proteins of the invention. Such instructions or package inserts may also address the particular advantages of the albumin fusion proteins of the inventions, such as the ease of storage for formulations that  
35 may require use in the field, outside of controlled hospital, clinic or office conditions. As described above, formulations of the invention may be in aqueous form and may be stored

under less than ideal circumstances without significant loss of therapeutic activity.

Albumin fusion proteins of the invention can also be included in nutraceuticals. For instance, certain albumin fusion proteins of the invention may be administered in natural products, including milk or milk product obtained from a transgenic mammal which expresses albumin fusion protein. Such compositions can also include plant or plant products obtained from a transgenic plant which expresses the albumin fusion protein. The albumin fusion protein can also be provided in powder or tablet form, with or without other known additives, carriers, fillers and diluents. Nutraceuticals are described in Scott Hegenhart, *Food Product Design*, Dec. 1993.

The invention also provides methods of treatment and/or prevention of diseases or disorders (such as, for example, any one or more of the diseases or disorders disclosed herein) by administration to a subject of an effective amount of an albumin fusion protein of the invention or a polynucleotide encoding an albumin fusion protein of the invention ("albumin fusion polynucleotide") in a pharmaceutically acceptable carrier.

The albumin fusion protein and/or polynucleotide will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the albumin fusion protein and/or polynucleotide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of the albumin fusion protein administered parenterally per dose will be in the range of about 1ug/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the albumin fusion protein is typically administered at a dose rate of about 1 ug/kg/hour to about 50 ug/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Albumin fusion proteins and/or polynucleotides can be administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular,

intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

Albumin fusion proteins and/or polynucleotides of the invention are also suitably administered by sustained-release systems. Examples of sustained-release albumin fusion proteins and/or polynucleotides are administered orally, rectally, parenterally, intracisternally, 5 intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, 10 subcutaneous and intraarticular injection and infusion. Additional examples of sustained-release albumin fusion proteins and/or polynucleotides include suitable polymeric materials (such as, for example, semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules), suitable hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, and sparingly soluble derivatives (such as, for 15 example, a sparingly soluble salt).

Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and Langer, Chem. Tech. 12:98-105 (1982)), 20 ethylene vinyl acetate (Langer et al., Id.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988).

Sustained-release albumin fusion proteins and/or polynucleotides also include liposomally entrapped albumin fusion proteins and/or polynucleotides of the invention (see generally, Langer, *Science* 249:1527-1533 (1990); Treat et al., in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 25 317 -327 and 353-365 (1989)). Liposomes containing the albumin fusion protein and/or polynucleotide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. (USA) 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci.(USA) 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. 30 Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal Therapeutic.

In yet an additional embodiment, the albumin fusion proteins and/or polynucleotides of the invention are delivered by way of a pump (see Langer, *supra*; Sefton, CRC Crit. Ref. 35 Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)).

Other controlled release systems are discussed in the review by Langer (*Science* 249:1527-1533 (1990)).

For parenteral administration, in one embodiment, the albumin fusion protein and/or polynucleotide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to the Therapeutic.

Generally, the formulations are prepared by contacting the albumin fusion protein and/or polynucleotide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The albumin fusion protein is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any pharmaceutical used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Albumin fusion proteins and/or polynucleotides generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.